Canonical representations of Gaussian processes and their applications

By

Takeyuki HIDA

Introduction

Let $B(t)$, $0 \leq t < \infty$, be an additive real Gaussian process with $E(B(t)) = 0$ and $F(t, u)$ be a real-valued function of (t, u). The process $\tilde{X}(t)$ defined as

$$(0.1) \qquad \tilde{X}(t) = \int_0^t F(t, u) dB(u)$$

is a real Gaussian process with mean 0, and enjoys the property

$$(0.2) \qquad \mathfrak{M}_t(\tilde{X}) \subset \mathfrak{M}_t(B), \qquad 0 \leq t < \infty,$$

where $\mathfrak{M}_t(\tilde{X})$ and $\mathfrak{M}_t(B)$ denote the closed linear manifolds generated by $\{\tilde{X}(\tau) \; ; \; \tau \leq t\}$ and $\{B(\tau) \; ; \; \tau \leq t\}$ respectively. Given a Gaussian process $X(t)$, P. Lévy called the expression (0.1) a *representation* of $X(t)$, if $\tilde{X}(t)$ is version of $X(t)$ and he introduced the concept of *canonical representation*. Roughly speaking, a canonical representation is one for which the equality holds instead of the inclusion relation in (0.2) (cf. Definition I. 2 and Theorem I. 2). In this case, $F(t, u)$ is called a canonical kernel.

P. Lévy has recently published several important papers concerning the canonical representation of Gaussian processes. However his pioneering works contain some points difficult for us to follow. The main aim of this paper is to establish his theory systematically and to prove some new facts.

We shall here give a brief account of the contents of this paper. In this paper we shall treat only real Gaussian processes and often omit the adjective "real Gaussian".

Section I. General theory.

As P. Lévy proved in [4],[1] a canonical representation of any process is uniquely determined if it exists. We shall prove this fact in detail in § I. 2. Further we shall give a necessary and sufficient condition for the existence of the canonical representation, using Hellinger-Hahn's Theorem in the theory of Hilbert Space. This fact is not found in Lévy's paper.

As to whether a given representation is canonical or not, Lévy gave a criterion using Hellinger integral (P. Lévy [6]). But we shall give another criterion which proves to be a generalization of Karhunen's kernel criterion for the moving average representation of stationary processes.

Section II. Multiple Markov process.

J. L. Doob (for stationary processes) and P. Lévy defined N-ple Markov processes[2] using the derivatives up to the $(N-1)$-th order of the processes. We generalize this notion to treat the processes which are not always differentiable.

The main results obtained here are as follows. The canonical kernel of the N-ple Markov process is a Goursat kernel of order N. This generalizes the fact obtained by Lévy. In § II. 3, we shall prove that the stationary N-ple Markov process is the sum of special simple Markov processes which are to be called general Ornstein-Uhlenbeck's Brownian motions. This also generalizes Doob's Theorem [1].

Section III. Lévy's $M(t)$ process.

Let $X(A)$, $A \in E^N$ (E^N is N-dimensional Euclidean space), be an ordinary Brownian motion with N-dimensional parameter (cf. P. Lévy [4]) and $M_N(t)$ be the average of $X(A)$ over the sphere with center O (origin of E^N) and radius $t(\geqq 0)$ in the parameter space E^N. $M_N(t)$ is clearly a Gaussian process with time parameter t. Lévy discussed the canonical representation and the multiple Markov property for this process $M_N(t)$ only in the case N is odd. We shall simplify his proof by transforming $M_N(t)$ into a stationary

---

1) Numbers in square brackets refer to the list of references at the end of the paper.

2) In Lévy's terminology "Markov process of order N in the restricted sense".

process. Our present method is applicable to the case N is even. We shall prove that $M_{2p}(t)$ is not a multiple Markov process but the limiting process of multiple Markov processes. Furthermore we shall prove that there are 2^{p-1} different representation of $M_{2p-1}(t)$, which provides an affirmative answer to Lévy's problem (Lévy [4, p. 146])

I would like to express my hearty thanks to Professor K. Itô for his encouragement and valuable suggestions and to Mr. N. Ikeda who helped me with valuable discussions in overcoming the difficulties in the course of this paper; in particular, the idea of using the reproducing kernel in the proof of Theorem I.4 is due to Mr. Ikeda.

Section I. General theory of representation.

§I.1. Definitions.

In order to define a representation of the given Gaussian process precisely, it is necessary for us to consider integrals with respect to certain random measures. The parameter space T of a process that will be treated here may be a closed interval, $(-\infty, \infty)$ or $[0, \infty)$. The symbol $\boldsymbol{B}_T$ denotes the Borel field of subsets of T.

Let $B(M)$, $M \in \boldsymbol{B}_T$, be a real Gaussian random measure such that

(I.1) $E(B(M)) = 0$ and $E(B(M)^2) = v(M)$ for every $M \in \boldsymbol{B}_T$,

where v is a (non-negative) measure defined on $\boldsymbol{B}_T$. Then, $B(M)$ can be decomposed into two parts in the following way

$$B(M) = B_1(M) + \sum_{t_j \in M} X_{t_j},$$

where $B_1(\cdot)$ is a random measure associated with the continuous measure $v_1(\cdot) = E(B_1(\cdot)^2)$ and X_{t_j}'s are mutually independent Gaussian random variables with mean 0, each one of which corresponds to the jump point t_j of $v(u) = v((-\infty, u] \cap T)$. $B_1(\cdot)$ and $\{X_{t_j}\}$ will be called the continuous part and discontinuous part of $B(\cdot)$ respectively.

Let $f(u)$ be a Borel measurable function. Then, if $f \in L^2(v)$, that is, $f \in \boldsymbol{L}^2(v_1)$ and $\sum_{t_j} f(t_j)^2 E(X_{t_j}^2) < \infty$, then

$$\int_M f(u)dB_1(u) \quad \text{and} \quad \sum_{t_j \in M} f(t_j)X_{t_j}$$

are well defined for any Borel set M. The integral of f over M may be written in the form

(I. 2) $$\int_M f(u)dB(u) = \int_M f(u)dB_1(u) + \sum_{t_j \in M} f(t_j)X_{t_j},$$

where the integrals which appear are interpreted in the usual way with respect to random measures dB, dB_1, Doob [2; pp. 426-433.]

Now we can give

Definition I. 1. Let $Y(t)$, $t \in T$, be a real Gaussian process with $E(Y(t))=0$ for every $t \in T$. Then the triple $(dB(t), \mathfrak{M}_t, F(t, u))$, or simply the pair $(dB(t), F(t, u))$, is called a *representation* of $Y(t)$, if

i) $B(\cdot)$ is a random measure satisfying (I. 1);

ii) $F(t, u)$ is a real Borel measurable function of u vanishing for $u>t$ and belonging to $L^2(v)$ for every t;

iii) $$X(t) = \int^t F(t, u)dB(u)^{3)}$$

is a version of $Y(t)$;

iv) $\mathfrak{M}_t$ is the closed linear manifold generated by $\{X(\tau) ; \tau \leq t\}$. The function $F(t, u)$ is called a *kernel* of the representation.

There are many examples of processes which have no representation. Furthermore, even if a process has a representation, it may not be uniquely determined. The following examples will serve to illustrate such circumstances.

Example I. 1. Let $Y_1(t)$ be a Gaussian process with covariance function $\Gamma(s, t)=1$ for $s=t$, and $=0$ for $s \neq t$, and with $E(Y_1(t))=0$. Then $Y_1(t)$ has no representation.

Example I. 2. Let $B_1(t)$ and $B_2(t)$, $0 \leq t < \infty$, be standard Brownian motions which are independent of each other. Define

$$Y_2(t) = \begin{cases} B_1(t) & \text{if } t \text{ is rational,} \\ B_2(t) & \text{if } t \text{ is irrational.} \end{cases}$$

Then $Y_2(t)$ has no representation. Detailed discussions concerning this will be given later.

---

3) The notation $\int^t \cdots$ means the integral $\int_{(-\infty, t] \cap T} \cdots$ in the sense of (I. 2).

Example I.3. Let $B(t)$, $0 \leq t < \infty$, be a standard Brownian motion. Then, for every positive integer n, we can determine constants $c_0, c_1, \cdots, c_n$ so that

$$\tilde{B}(t) = \int_0^t \left(c_0 + c_1 \frac{u}{t} + c_2 \left(\frac{u}{t} \right)^2 + \cdots + c_n \left(\frac{u}{t} \right)^n \right) dB(u)$$

is again a standard Brownian motion (P. Lévy [6]). This proves that $B(t)$ has infinitely many representations.

We have now to determine the best class of representations for our purpose among all posible representations of a given process.

Definition. I.2. The representation $(dB(t),\ \mathfrak{M}_t,\ F(t, u))$ is called *canonical*, if

$$E(X(t)/\boldsymbol{B}_s) = \int^s F(t, u) dB(u)$$

holds for every $s \leq t$, where $\boldsymbol{B}_s$ is the smallest Borel field of measurable ω-sets with respect to which all the $X(\tau)$'s $(\tau \leq s)$ are measurable. In this case, $F(t, u)$ is called a *canonical kernel*.

Definition. I.3. Two representations $(dB^{(i)}(t),\ \mathfrak{M}_t^{(i)},\ F^{(i)}(t, u))$, $i = 1, 2$, are called *equivalent*, if

$$\int_M F^{(1)}(t, u)^2 dv^{(1)}(u) = \int_M F^{(2)}(t, u)^2 dv^{(2)}(u) \qquad \text{for every } M \in \boldsymbol{B}_T,$$

considering them as measures, where

$$dv^{(i)}(u) = E(dB^{(i)}(u)^2), \qquad i = 1, 2.$$

This relation obviously satisfies the equivalence relations and therefore we can get the classes of representations.

Theorem I.1. (*P. Lévy*) *For every* $Y(t)$, *there exists at most one class of canonical representations.*

Proof. Let $(dB^{(i)}(t),\ \mathfrak{M}_t^{(i)},\ F^{(i)}(t, u))$ $i = 1, 2$, be canonical representations of $Y(t)$. Writing

$$X^{(i)}(t) = \int^t F^{(i)}(t, u) dB^{(i)}(u), \qquad i = 1, 2,$$

we have, for every t and every $s(\leq t)$,

$$E((X^{(i)}(t)/\boldsymbol{B}_s^{(i)}) = \int^s F^{(i)}(t, u) dB^{(i)}(u),$$

where $\boldsymbol{B}_t^{(i)}$ denotes the Borel field corresponding to $X^{(i)}(t)$.

The equality

$$E(E(X^{(1)}(t)/\boldsymbol{B}_s^{(1)})^2) = E(E(X^{(2)}(t)/\boldsymbol{B}_s^{(2)})^2)$$

holds, since both sides are determined only by the probability law of $X(t)$. Hence we have

$$\int^s F^{(1)}(t, u)^2 dv^{(1)}(u) = \int^s F^{(2)}(t, u)^2 dv^{(2)}(u), \qquad \text{for every } s(\leq t),$$

which proves the theorem.

§I. 2. Canonical representations.

In this article we shall study important properties of a canonical representation.

Definition I. 4.   A canonical representation $(dB(t), \mathfrak{M}_t, F(t, u))$ is called *proper* if

(I. 3) $\mathfrak{M}_t = \mathfrak{M}_t(B)$ for every $t \in T$,

where $\mathfrak{M}_t(B)$ is the closed linear manifold generated by

$$\left\{ \int_{(-\infty, t] \cap M} dB(u) \; ; \; M \in \boldsymbol{B}_T \right\}.$$

Theorem I. 2.   *For any given canonical representation, we can construct an equivalent proper canonical representation.*

Proof.   It is sufficient to consider the case in which v is a continuous measure.   Let $(dB(t), \mathfrak{M}_t, \widetilde{F}(t, u))$ be a given canonical representation of $Y(t)$.   We shall show that we can construct a proper canonical representation of $Y(t)$ by deforming the given one.

1°) *Deformation.*   First define $F(t, u) = \widetilde{F}(t, u)$.   Put

$$\mu(M) = \bigvee_t \left(\int_M F(t, u)^2 dv(u) \right),^{4)}$$

$$v(M) \equiv \int_M dv(u) = E\left(\int_M dB(u) \right)^2, \qquad M \in \boldsymbol{B}_T.$$

Here we may suppose that v is a continuous measure.   Then

$$\mu \ll v ;^{5)}$$

hence, by Radon-Nykodym's Theorem, there exists a Borel measurable function $f(u) \geq 0$ such that

---

4) $\bigvee$ means the lattice sum.

5) $\mu \ll \widetilde{v}$ means that the measure μ is absolutely continuous with respect to the measure $\widetilde{v}$.

$$\mu(M) = \int_M f(u)dv(u) .$$

Since $N = N(f) = \{u ; f(u) > 0\}$ is Borel measurable, we can construct a random measure

(I. 4) $$B(M) \equiv \int_M dB(u) = \int_M \chi_N(u)d\tilde{B}(u), \qquad M \in \boldsymbol{B}_T ,$$

where $\chi_N(u)$ is the indicator function of N.

2°) The triple $(dB(t), \mathfrak{M}_t, F(t, u))$ is a representation of $Y(t)$. To prove this, it is sufficient to show that $X(t) = \int^t F(t, u)dB(u)$ is the same process as $Y(t)$, in the sense of equivalence in law. This can be proved as follows.

(I. 5)
$$\begin{aligned}
&E\left(\int^t \tilde{F}(t, u)d\tilde{B}(u) - \int^t F(t, u)dB(u)\right)^2 \\
&= E\left(\int^t F(t, u)(1 - \chi_N(u))d\tilde{B}(u)\right)^2 \\
&= \int^t F(t, u)^2(1 - \chi_N(u))^2 d\tilde{v}(u) \\
&= \int_{(-\infty, t] \cap N} (1 - \chi_N(u))^2 F(t, u)^2 d\tilde{v}(u) = 0 ,
\end{aligned}$$

which shows that $X(t) = \tilde{X}(t) \left(= \int^t \tilde{F}(t, u)d\tilde{B}(u)\right)$ with probability one. Since $\tilde{X}(t)$ is the same process as $Y(t)$ by assumption, the relation above proves our assertion.

3°) Finally we shall prove that the representation $(dB(t), \mathfrak{M}_t, F(t, u))$ is proper canonical. Now we prove

(I. 6) $$\mathfrak{M}_t \supset \mathfrak{M}_t(B) .$$

Suppose an element Z of $\mathfrak{M}_t(B)$ is orthogonal to (hence independent of) $\mathfrak{M}_t$, that is, $E(Z \cdot X(s)) = 0$ for every $s \leq t$. Then Z is orthogonal to $E(X(s'')/\boldsymbol{B}_{s'})$ for every $s' \leq t$ and every s'', since it is an element of $\mathfrak{M}_t$.

On the other hand Z can be written as

$$Z = \int^t h(u)dB(u) ,$$

with a Borel measurable function h. Noting that

$E(X(s'')/\boldsymbol{B}_{s'}) = $ Projection of $X(s'')$ on $\mathfrak{M}_{s'}$ (in the $\boldsymbol{L}^2$ sense)

$$= \text{Projection of } \tilde{X}(s'') \text{ on } \mathfrak{M}_{s'}$$

(since $X(t) = \tilde{X}(t)$ with probability 1)

$$= E(\tilde{X}(s'')/\boldsymbol{B}_{s'}) = \int^{s'} \tilde{F}(s'', u)d\tilde{B}(u)$$

(from canonical property)

$$= \int^{s'} F(s'', u)dB(u) ,$$

we have

$$E(Z \cdot E(X(s''))/\boldsymbol{B}_{s'})) = \int^{s'} h(u)F(s'', u)d\bar{v}(u) = 0 .$$

Therefore, for every $s, s'(<t)$

$$\int_s^{s'} h(u)F(s'', u)d\bar{v}(u) = 0 .$$

Since s'' is arbitrary, we can prove

$$\mu(N(h)) = 0 ,$$

where $N(h) = \{u ; h(u) \neq 0\}$. Hence we have

$$E(Z \cdot W) = 0 \quad \text{for every } W \in \mathfrak{M}_t(B) ,$$

Thus we have proved (I. 6). Consequently we have (I. 3).
    Now from (I. 3),

$$E(X(t)/B_s) = \text{projection of } X(t) \text{ on } \mathfrak{M}_s$$

$$= \text{Projection of } X(t) \text{ on } \mathfrak{M}_s(B) = \int^s F(t, u)dB(u) ,$$

which proves that the representation is canonical and (I. 6) implies that it is proper. Thus we have proved the theorem.
    By the argument used in the proof of the theorem, we hav

Corollary. If a representation $(dB(t), \mathfrak{M}_t, F(t, u))$ (not necessarily canonical) satisfies the condition

$$\mathfrak{M}_t(B) = \mathfrak{M}_t \quad \text{for every } t,$$

then it is (proper) canonical.
    As is well known, a stationary process $X(t)$ which is purely non-deterministic and M_2-continuous can be expressed as

$$X(t) = \int_{-\infty}^t F(t-u)dB(u)$$

and there exists one and only one representation having the property (I. 3). (Karhunen [1]). In our case, in which $X(t)$ is Gaussian, this means that it has a proper canonical representation.

§ I. 3. Existence of representation.

In order to study the existence of the canonical representation we shall summarize some known theorems of the theory of Hilbert space.

Let $\Gamma(s, t)$, $s, t \in T$, be a real non-negative definite function. Then there exists a Hilbert space $\mathfrak{H}$ satisfying

  i) $\Gamma(s, t)$ belongs to $\mathfrak{H}$ as a function of s,

  ii) $\langle f(s), \Gamma(s, t) \rangle^{6)} = f(t)$ for every $f \in \mathfrak{H}$,

  iii) $\mathfrak{H}$ is the closed (in the topology $\| \; \|$) linear manifold generated by $\{\Gamma(\cdot, t) ; t \in T\}$.

$\Gamma(s, t)$ is the *reproducing kernel* of that Hilbert space.

The construction and the important properties of $\mathfrak{H}$ may be seen in Aronszajn [1].

We can construct sub-spaces $\mathfrak{H}_t$ and $\mathfrak{H}_t^*$ of $\mathfrak{H}$:

$$\mathfrak{H}_t = \text{sub-space of } \mathfrak{H} \text{ generated by } \{\Gamma(\cdot, \tau) ; \tau \leq t\} ;$$

$$\mathfrak{H}_t^* = \bigcap_n \mathfrak{H}_{t+\frac{1}{n}} .$$

Now let us assume that

(H. 1) $\mathfrak{H}$ is separable,

(H. 2) $\bigcap_{t \in T} \mathfrak{H}_t = \{0\}$ (hence $\bigcap_{t \in T} \mathfrak{H}_t^* = \{0\}$) .

Noting that

$$\bigvee_{t \in T} \mathfrak{H}_t^* = \mathfrak{H} \quad \text{and} \quad \mathfrak{H}_s^* \subset \mathfrak{H}_t^*, \qquad s \leq t ,$$

we can see that there exists a resolution of the identity $\{E(t) ; t \in T\}$ such that

(I. 7) $\mathfrak{H}_t^* = E(t)\mathfrak{H} ,$

by assumption. Then, by Hellinger-Hahn's Theorem,[7] there exist two denumerable sets $\{f^{(i)}\}$, $i = 1, 2, \cdots$ and $g^{(j)l}$, $j, l = 1, 2, \cdots$ in $\mathfrak{H}$ satisfying the following conditions (I. 8) to (I. 10).

---

6) The symbol $\langle \; , \; \rangle$ denotes the inner product. We shall use $\| \; \|$ to denote the norm, i.e. $\| f \| = \sqrt{\langle f, f \rangle}$.

7) For proof see M. H. Stone [1] or S. Itô [1].

$$(I. 8) \begin{cases} \text{i)} & \text{For any intervals } \Delta_1, \Delta_2 \\ & \langle \Delta_1 E f^{(i)}, \Delta_2 E f^{(j)} \rangle = 0, \quad i \neq j, \\ & \text{with } \Delta E = E(b) - E(a) \text{ for } \Delta = (a, b]; \\ \text{ii)} & \text{if } \Delta_1 \cap \Delta_2 = \emptyset \\ & \langle \Delta_1 E f^{(i)}, \Delta_2 E f^{(i)} \rangle = 0; \\ \text{iii)} & \text{for any } i, \ \rho_i(t) = \|E(t) f^{(i)}\| \text{ is continuous, non-decreasing} \\ & \text{and } \rho_{i+1} \ll \rho_i \text{ (considered as measures)}; \\ \text{iv)} & g^{(j)l} \text{ is the eigenvector of the self-adjoint operator } H = \\ & \int t dE(t) \text{ corresponding to the eigenvalue } t_j (l = 1, 2, \cdots). \end{cases}$$

$$(I. 9) \qquad \mathfrak{H} = \mathfrak{M} \oplus \mathfrak{N} \qquad \text{(direct sum)},$$

where $\mathfrak{M}$ and $\mathfrak{N}$ are defined by

$$\mathfrak{M} = \sum_i \oplus \mathfrak{M}(f^{(i)}),$$

$$\mathfrak{M}(f^{(i)}) = \left\{ f; f = \int \varphi(t) dE(t) f^{(i)}, \qquad \varphi \in L^2(\rho_i) \right\};$$

$$\mathfrak{N} = \sum_{t_j} \sum_l \oplus \mathfrak{N}(g^{(j)l}),$$

$$\mathfrak{N}(g^{(j)l}) = \text{one dimensional sub-space generated by } g^{(j)l};$$

$$(I. 10) \qquad E(t) \mathfrak{M}(f^{(i)}) \subset \mathfrak{M}(f^{(i)}),$$

which is equivalent to

$$(I. 10') \qquad E(t) P_i = P_i E(t), \qquad P_i = \text{Projection on } \mathfrak{M}(f^{(i)}).$$

Furthermore, though there may be many ways of choosing such $\{f^{(i)}\}$ and $\{g^{(j)l}\}$, their numbers are always the same.

By virtue of this theorem, we can define the multiplicity of $E(t)$.

Definition I. 5. The supremum of the number of $f^{(i)}$'s and the numbers of linearly independent eigenvectors corresponding to each t_j is called the *multiplicity* of $E(t)$.

Theorem I. 3. $\Gamma(\cdot, t)$ *is expressible as*

$$(I. 11) \qquad \Gamma(\cdot, t) = \sum_i \int^t F_i(t, u) dE(u) f^{(i)} + \sum_{t_j \leq t} \sum_l b_j^l(t) g^{(j)l}.$$

Proof. From (I. 9), $\Gamma(\cdot, t)$ is written in the form

$$\Gamma(\cdot, t) = \sum_i \int F_i(t, u) dE(u) f^{(i)} + \sum_{t_j} \sum_l b_j^l(t) g^{(j)l}.$$

Applying (I. 10), we have

$$\Gamma(\cdot, t) = E(t)\Gamma(\cdot, t) = \sum_i E(t) \int F_i(t, u) dE(u) f^{(i)} + \sum_{t_j} \sum_l b_j^l(t) E(t) g^{(j)l}$$

$$= \sum_i \int^t F_i(t, u) dE(u) f^{(i)} + \sum_{t_j \leq t} \sum_l b_j^l(t) E(t) g^{(j)l} ,$$

where the summation in the second term in the equation above extends over those $g^{(j)l}$'s the eigenvalue of which are not larger than t.

The case in which the function $F_i(t, u)$ in (I. 11) is degenerate, for example,

(I. 12) $$\sum_{k=1}^N f_k(t) g_k(u) ,$$

is of special interest, as we shall see in the next section.

After the preparation above, we can now discuss the existence of the canonical representation. Given a process $Y(t)$, let $\Gamma(s, t)$ be its covariance function. Let $\mathfrak{M}_t$ be the closed linear manifold generated by $\{Y(\tau) ; \tau \leq t\}$ and

$$\mathfrak{M} = \bigvee_{t \in T} \mathfrak{M}_t, \quad \mathfrak{M}_t^* = \bigwedge_n \mathfrak{M}_{t+1/n} .$$

We shall assume that

($\mathfrak{M}$. 1) $\mathfrak{M}$ is separable (as a sub-space of $L^2(\Omega)$),

($\mathfrak{M}$. 2) $\bigwedge_{t \in T} \mathfrak{M}_t = \{0\}$.

We shall prove the following preliminary theorem leading to the fundamental Theorems I. 5 and I. 6.

Theorem I. 4.   *There exists an isometric transformation from $\mathfrak{H}$ onto $\mathfrak{M}$ defined by*

(I. 13) $$\mathfrak{H} \ni \Gamma(\cdot, t) \leftrightarrow Y(t) \in \mathfrak{M} .$$

This isometry induces the following correspondence:
i) $\mathfrak{H}_t \leftrightarrow \mathfrak{M}_t, \quad \mathfrak{H}_t^* \leftrightarrow \mathfrak{M}_t^*,$
ii) $E(t)f \leftrightarrow E(X/\boldsymbol{B}_t^*)$ *provided that* $f \leftrightarrow X$, *with* $\boldsymbol{B}_t^* = \boldsymbol{B}(\mathfrak{M}_t^*)$.

Proof.   Define a mapping S from $\boldsymbol{L} = \{\Gamma(\cdot, t) ; t \in T\}$[8] into $\mathfrak{M}$ by

---

8) $\{ \cdots \}$ denotes the linear space generated by the elements that are written in the bracket.

$$S: \quad \Gamma(\cdot, t) \rightarrow Y(t)$$
$$\sum a_i \Gamma(\cdot, t_i) \rightarrow \sum a_i Y(t_i) \quad (a_i : \text{ real})$$

Then S is a linear transformation from a linear space L into the linear space $\mathfrak{L} = \{Y(t), \ t \in T\} \subset \mathfrak{M}$. Suppose $\sum a_i Y(t_i) = 0$. Then $\sum a_i \Gamma(t, t_i) = 0$ for every $t \in T$, which implies

$$f(t) = \langle f(\cdot), \ \Gamma(\cdot, t) \rangle = 0, \quad \text{if } f(\cdot) = \sum a_i \Gamma(\cdot, t).$$

This shows that S is a one-to-one mapping from L onto $\mathfrak{L}$. And further

$$\langle \Gamma(\cdot, t), \ \Gamma(\cdot, s) \rangle = \Gamma(s, t) = E(Y(t) \cdot Y(s)),$$

which proves that S is isometric.

Since L and $\mathfrak{L}$ are dense in $\mathfrak{H}$ and $\mathfrak{M}$ respectively, S can be extended to an isometric one-to-one linear transformation $\bar{S}$ from $\mathfrak{H}$ onto $\mathfrak{M}$. Hence we have proved the existence of the isometry.

Next, $\bar{S}\mathfrak{H}_t = \mathfrak{M}_t$ is obvious. Therefore if $f \leftrightarrow X$,

$$E(t)f \ \leftrightarrow \ \text{Projection of } X \text{ on } \mathfrak{M}_t^* = E(X/\boldsymbol{B}_t^*).$$

Thus we have proved ii).

Theorem I.4 along with the assumptions $(\mathfrak{M}.1)$, $(\mathfrak{M}.2)$ implies that $(H.1)$ and $(H.2)$ hold. We can therefore appeal to the Hellinger-Hahn's Theorem. Let $f^{(i)}$ be as defined in the statement of that theorem. Then there exists a continuous additive process $B^{(i)}(t)$ such that

$$dE(t)f^{(i)} \ \leftrightarrow \ dB^{(i)}(t)$$

under the correspondence mentioned in Theorem I.4.

Theorem I.5. *If $Y(t)$ satisfies $(\mathfrak{M}.1)$ and $(\mathfrak{M}.2)$, there exist Gaussian random measures $\{B^{(i)}(\cdot)\}$ and random variables $Y^l_{t_j}$ such that*

i) $B^{(i)}(\cdot), \ Y^l_{t_j}, \quad i, j, l = 1, 2, \cdots$ *are all independent,*

ii) $E(B^{(i+1)}(\cdot)^2) \ll E(B^{(i)}(\cdot)^2), \quad i = 1, 2, \cdots,$

iii) $Y(t) = \sum_i \int^t F_i(t, u) dB^{(i)}(u) + \sum_{t_j \leq t} \sum_l b^l_j(t) Y^l_{t_j},$

iv) $E(Y(t)/\boldsymbol{B}_s) = \sum_i \int^s F(t, u) dB^{(i)}(u) + \sum^*_{t_j \leq s} \sum_l b^l_j(t) Y^l_{t_j},$[9] $s \leq t.$

---

9) $\displaystyle \sum^*_{t_j \leq s} \sum_l = \sum_{t_j < s} \sum_l + \sum_{l(b^l_j(s) \neq 0)}$

Proof. If follows from (I. 8), (I. 9) and Theorem I. 4 that i), ii) and iii) are satisfied. Considering the isometry $\bar{S}$ in the proof of Theorem I. 4, we have

$$E(Y(t)/\boldsymbol{B}_s^*) = \bar{S}^{-1}(E(s)\Gamma(\cdot, t))$$

$$= \bar{S}^{-1}(E(s) \sum_i \int^t F_i(t, u)dE(u)f^{(i)} + E(s) \sum_{t_j \leq t} \sum_l b_j^l(t)g^{(j)l}) .$$

By (I. 10), this is equal to

$$\bar{S}^{-1}(\sum_i \int^s F_i(t, u)dE(u)f^{(i)} + \sum_{t_j \leq s}^* \sum_l b_j^l(t)g^{(j)l})$$

$$= \sum_i \int^s F_i(t, u)dB^{(i)}(u) + \sum_{t_j \leq s}^* \sum_l b_j^l(t)Y_{t_j}^l$$

$$(E(X(t)/\boldsymbol{B}_s) = E(E(X(t)/\boldsymbol{B}_s^*)/\boldsymbol{B}_s)) ,$$

which proves iv).

Definition I. 6. The system $(dB^{(i)}(t), \mathfrak{M}_t^{(i)}, F(t, u), b_j^l(t) Y_{t_j}^l; i, j, l = 1, 2, \cdots)$ obtained above is called a *generalized canonical representation* of $Y(t)$. The multiplicity of $E(t)$ is called the *multiplicity* of $Y(t)$.

The classification of generalized canonical representations is very complicated, because, for one thing, the choice of the system $\{f^{(i)}\}$ is not unique.

Theorem I. 6. *A necessary and sufficient condition that $Y(t)$ has a canonical representation is that $Y(t)$ satisfies the conditions* $(\mathfrak{M}. 1)$, $(\mathfrak{M}. 2)$ *and*

$(\mathfrak{M}. 3)$ *The multiplicity of $Y(t)$ is one.*

Proof. Necessity. Suppose that $(dB(t), \mathfrak{M}_t, F(t, u))$ is a canonical representation of $Y(t)$ and

(I. 14) $X(t) = \int^t F(t, u)dB(u) \sim Y(t),$ $(\sim:$ equivalent in law$)$

The separability of $\mathfrak{M} = \bigvee_{t \in T} \mathfrak{M}_t$ is easily deduced from the definition of the integral.

Let $\mathfrak{M}_t(B)$ be the same as in Definition I. 4. Then

$$\mathfrak{M}_t \subset \mathfrak{M}_t(B) = \left\{ \int^t f(u)dB(u) ; f \text{ is Borel measurable and } \in L^2(v) \right\}$$

and

$$\bigwedge_{t \in T} \mathfrak{M}_t(B) = \{0\}$$

imply the condition ($\mathfrak{M}$.2). ($\mathfrak{M}$.1) follows at once from the separability of $\bigvee_{t \in T} \mathfrak{M}_t(B)$.

Finally, if we decompose the random measure $B(\cdot)$ into two parts in the same way as in (I.1), we can easily see that the multiplicity of $Y(t)$ is one by Definition I.6.

Sufficiency. From Theorem I.5 and the assumption that the multiplicity of $Y(t)$ is one, $Y(t)$ can be expressed in the form

$$Y(t) = \int^t F_1(t, u) dB^{(1)}(u) + \sum_{t_j \leqq t} b_j(t) X_{t_j}.$$

Now we can define a random measure $B(\cdot)$ by

$$B((a, b]) \equiv B(b) - B(a) = \int_a^b dB^{(1)}(u) + \sum_{a < t_j \leqq b} a_j X_{t_j}, \quad (\sum_j a_j^2 E(X_{t_j}^2) < \infty).$$

And define a function $F(t, u)$ of u for every fixed t, by

$$F(t, u) = \begin{cases} F_1(t, u) & \text{if } u \in T \bigcap (-\infty, t] - \{t_j\}, \\ b_j(t)/a_j & \text{if } u = t_j. \end{cases}$$

Then we have

$$\int^t F(t, u) dB(u) \equiv \int^t F(t, u) dB^{(1)}(u) + \sum_{t_j \leqq t} a_j F(t, t_j) X_{t_j}.$$

$$= \int^t F_1(t, u) dB^{(1)}(u) + \int^t (F(t, u) - F_1(t, u)) dB^{(1)}(u) + \sum_{t_j \leqq t} b_j(t) X_{t_j},$$

which is equal to $Y(t)$, since the second term of the last expression is zero with probability one.

The canonical property of $(dB(t), \mathfrak{M}_t, F(t, u))$ follows from iv) in Theorem I.5.

The process $Y_2(t)$ which was given in Example I.2 has no canonical representation, because the multiplicity of $Y_2(t)$ is 2, as is easily seen. But a generalized canonical representation exists. In fact it can be expressed in the form

$$Y_2(t) = \int_0^t F(t, u) dB_1(u) + \int_0^t (1 - F(t, u)) dB_2(u),$$

where

$$F(t, u) = \begin{cases} 1, & \text{if } t \text{ is rational}, \\ 0, & \text{if } t \text{ is irrational}. \end{cases}$$

Corollary. *An additive process has a canonical representation.*

Proof. If $B(t)$ is an additive process, it can be expressed in the form

$$B(t) = \int^t dB_1(t) + \sum_{t_j \leq t}^{*} X_{t_j}.$$

This shows that the multiplicity is one, and our assertion follows.

§I. 4. Kernel criterion for canonical representation.

It is important to give a criterion to determine whether a given representation is canonical or not. P. Lévy gave a criterion involving a Hellinger's integral (Lévy [6]), but we shall give another.

By Theorem I. 2, it is sufficient to give a criterion for a proper canonical representation.

Theorem I. 7. *A representation* $(dB(t), \mathfrak{M}_t, F(t,u))$ *is proper canonical if and only if, for any fixed* $t_0 \in T$

(I. 15) $\displaystyle\int^t F(t, u) f(u) dv(u) = 0$ *for every* $t \leq t_0$.

implies

(I. 16) $f(u) = 0$ *almost everywhere* (v) *on* $(-\infty, t_0] \bigcap T$.

Proof. Suppose that the given representation is not proper canonical. Then by (I. 3) there exists an element $Z(\neq 0)$ of $\mathfrak{M}_{t_0}(B)$ which is independent of every $X(t)$, $t \leq t_0$. Noting that Z can be expressed in the form

$$Z = \int^{t_0} f(u) dB(u), \qquad f \in \boldsymbol{L}^2(v),$$

we have

$$E(Z \cdot X(t)) = 0, \qquad \text{for every } t \leq t_0,$$

which is identical with (I. 15). On the other hand

$$0 \neq E(Z^2) = \int^{t_0} f(u)^2 dv(u).$$

This shows that (I. 16) does not hold.

Conversely, if there is a function $f(u)$ satisfying (I. 15) but not satisfying (I. 16) for some t_0, then

$$Z = \int^{t_0} f(u) dB(u)$$

belongs to $\mathfrak{M}_t(B)$ but does not belong to $\mathfrak{M}_t$. Hence the representation is not proper canonical.

According to Karhunen [1], a stationary process which is purely non-deterministic and M_2-continuous always has its (moving average) representation, the kernel of which is a function of $t-u$. He also gave a kernel criterion for canonical (in our terminology) representation. One can easily see that our Theorem I. 7. is a generalization of Karhunen's theorem

Example 1. 4. Let $X_1(t)$ and $X_2(t)$ be defined by

$$X_1(t) = \int_0^t (2t-u)dB_1(u)$$

$$X_2(t) = \int_0^t (-3t+4u)dB_2(u)$$

where $B_i(t)$, $i=1, 2$, are ordinary Brownian motions. Then the two processes have the same probability distribution (cf. P. Lévy [4]), since they have the common covariance function $3ts-2s^2/3$ $(t>s)$. Using Theorem I. 7, we can prove that $(dB_1(t), 2t-u)$ is a proper canonical representation of $X_1(t)$. On the other hand

$$Z = \int_0^{t_0} u dB_2(u)$$

is independent of every $X_2(t)$ $(t \leq t_0)$, which proves that the representation $(dB_2(t), -3t+4u)$ of $X_2(t)$ is not proper canonical— indeed it is not canonical.

Example 1. 5. (Particular case of Example I. 3). If we denote an ordinary Brownian motion by $B_0(t)$,

$$X(t) = \int_0^t (3-12u/t+10u^2/t^2)dB_0(u)$$

is again a Brownian motion. Here

$$Z_1 = \int_0^{t_0} u dB_0(u) \quad \text{and} \quad Z_2 = \int_0^{t_0} u^2 dB_0(u)$$

are independent of every $X(t)$ $(t \leq t_0)$. Hence $(dB_0(t), 3-12u/t +10u^2/t^2)$ is not proper canonical. In fact, the canonical representation of $B_0(t)$ is $(dB_0(t), 1)$.

Section II. Multiple Markov Gaussian Processes.

§ II. 1. Simple Markov Gaussian Processes.

We intend to study multiple Markov Gaussian processes in this section, using the general theory of representation. All the processes to be discussed here are Gaussian processes with mean 0 satisfying the conditions ($\mathfrak{M}$. 1) and ($\mathfrak{M}$. 2). Furthermore we may assume that

($\mathfrak{M}$. 4) $\qquad\qquad \mathfrak{M}_t$ is continuous in t,

that is,

$$\lim_{t \to t_0} \mathfrak{M}_t \quad \text{exists and is equal to} \quad \mathfrak{M}_{t_0},$$

since we can easily remove the discontinuity of $\mathfrak{M}_t$.

First we shall treat a simple Markov Gaussian process. Though some of the results are well known, our presentation of the results will stress their specific probabilistic significance form our standpoint.

Let $Y(t)$ be a simple Markov process. As $Y(t)$ is Gaussian the simple Markov property is equivalent to the condition that, if $s \leq t$,

(II. 1) $\qquad\qquad E(Y(t)/\boldsymbol{B}_s) = \varphi(t, s)Y(s)$,

where $\varphi(t, s)$ is a real valued ordinary function of (t, s) (Doob [2]). This is also equivalent to

(II. 2) $\qquad Y(t) - \varphi(t, s) Y(s)$ is independent of every $Y(\tau)$, $\qquad \tau \leq s$.

To avoid the case in which $Y(t)$ and $Y(s)$ are independent for $s \neq t$, let us assume that

(II. 3) $\qquad\qquad \Gamma(s, t) = E(Y(t) \cdot Y(s))$ never vanishes.

Then the equality

$$E(E(Y(t)/\boldsymbol{B}_{s'})/\boldsymbol{B}_s) = E(Y(t)/\boldsymbol{B}_s), \qquad \text{for every } s \leq s' \leq t$$

implies

$$\varphi(t, s')\varphi(s', s)Y(s) = \varphi(t, s)Y(s).$$

Since $\Gamma(s, s) \neq 0$ by the assumption (II. 3), we have

(II. 4) $\qquad\qquad \begin{cases} \varphi(t, s')\varphi(s', s) = \varphi(t, s), \\ \qquad \varphi(t, t) = 1, \end{cases}$

and we can prove that $\varphi(t, s)$ never vanishes. If we use the convention

$$\varphi(t, s) = \varphi(s, t)^{-1} \quad \text{for } s > t,$$

$\varphi(t, s)$ may be written as

$$\varphi(t, s) = f(t)/f(s),$$

where $f(t) = \varphi(t, s_0)$ with some fixed s_0.
Hence we have, from (II. 1),

$$E(f(t)^{-1}Y(t)/\boldsymbol{B}_s) = f(s)^{-1}Y(s),$$

which proves that $U(t) = f(t)^{-1}Y(t)$ is an additive process. Here we should note that the system of Borel fields relative to $U(t)$ is the same as that relative to $Y(t)$, since $f(t)$ never vanishes.

According to the Corollary to Theorem I. 6, $U(t)$ has a canonical representation, which has no discontinuous part as we assume ($\mathfrak{M}. 4$). Hence so does $Y(t)$:

$$Y(t) = f(t)U(t) = f(t)\int^t dU(u).$$

Conversely, a process expressed in this form is obviously a simple Markov process provided that $f(t)$ never vanishes.

Summing up, we have

Theorem II. 1. *Under the assumption* ($\mathfrak{M}. 1$), ($\mathfrak{M}. 2$), ($\mathfrak{M}. 4$) *and* (II. 3), *a necessry and sufficient condition that $Y(t)$ is a simple Markov process is that it can be expressed in the form*

$$\text{(II. 5)} \qquad Y(t) = f(t)U(t) = f(t)\int^t dU(u) = \int^t f(t)g(u)dB(u),$$

where $U(t)$ is an additive process with the property ($\mathfrak{M}. 4$) *($dB(t)$ is a continuous random measure) and $f(t)$ never vanishes.*

Making use of this theorem, we have (under the same assumptions)

Corollary 1. *Let $Y(t)$ be expressed in the form* (II. 5). *If $Y(t)$ is continuous in the mean, then $f(t)$ is continuous and $U(t)$ is continuous in the mean.*

Proof. If $Y(t)$ is continuous in the mean, then

$$\lim_{t \to t_0} E(Y(t)Y(s)) = E(Y(t_0)Y(s))$$

by the continuity of inner product in $\boldsymbol{L}^2(\Omega)$. By (II. 5), this can be written as,

$$\lim_{t \to t_0} f(t)f(s)E(U(t)U(s)) = f(t_0)f(s)E(U(s)^2), \qquad t > s,$$

since $U(t)$ is additive. Noting that $f(s) \neq 0$ and $E(U(s)^2) \neq 0$, we can see the continuity of $f(t)$. The continuity of $U(t)$ follows from $E(U(t)U(s)) = \Gamma(t, s)/(f(t)f(s))$.

In particular, if $T = [0, \infty)$ and $E(U(t)^2)$ has a continuous derivative, $U(t)$ becomes an ordinary Brownian motion by the change of time scale.[1] In other words, $Y(t)$ has a canonical representation $(dB_0(t), f(t)\sigma(u))$ with Wiener's random measure $B_0(\cdot)$ and a proper canonical kernel $f(t)\sigma(u)$.

Corollary 2. *If* $Y(t)$ *is a stationary simple Markov process satisfying the conditions* ($\mathfrak{M}$. 1), ($\mathfrak{M}$. 2) *and* (II. 3), *then it has a version*

$$(\text{II. 6}) \qquad c\int_{-\infty}^{t} e^{-\lambda(t-u)}dB_0(u), \qquad \lambda > 0.$$

Proof. As is easily seen in the proof of Theorem II. 1, the covariance function γ of $Y(t)$ can be written in the form

$$(\text{II. 7}) \qquad \gamma(h) = f(t+h)f(t)\sigma(t)^2, \qquad h > 0,$$

even though we do not assume ($\mathfrak{M}$. 4). Putting $t = 0$ in (II. 7), we have

$$f(h) = c_1\gamma(h),$$

and putting $h = 0$ in (II. 7), we have

$$\sigma(t)^2 = c_2f(t)^{-2} = c_3\gamma(h)^{-2}.$$

Hence it follows from (II. 7) that

$$c_3\gamma(h+t) = c_4\gamma(h) \cdot c_4\gamma(t).$$

Since γ is bounded above, we have

$$\gamma(h) = c^2e^{-\lambda|h|}, \qquad \lambda > 0.$$

§II. 2. Multiple Markov Gaussian processes.

In this article we shall define N-ple Markov process as a generalization of simple Markov process and study its properties. The property (II. 2) for a simple Markov process suggests that it is natural to give the following

Definition II. 1. If $\{E(Y(t_i)/\boldsymbol{B}_{t_0})\}$, $i = 1, 2, \cdots, N$, are linearly

---

1) See Seguchi-Ikeda [1].

independent for any $\{t_i\}$ with $t_0 \leq t_1 < t_2 < \cdots < t_N$, and if $\{E(Y(t_i)/B_{t_0})\}$, $i=1, 2, \cdots, N, N+1$, are linearly dependent for any $\{t_i\}$ with $t_1 < t_2 < \cdots < t_{N+1}$, then $Y(t)$ is called *N-ple Markov process.*

A simple Markov process is a 1-ple Markov process in this sense only if it satisfies (II. 3).

Theorem II. 2. *If $Y(t)$ is an N-ple Markov process satisfying* ($\mathfrak{M}$. 1), ($\mathfrak{M}$. 2), ($\mathfrak{M}$. 3) *and* ($\mathfrak{M}$. 4), *it has a version $X(t)$ expressed in the following form*

(II. 8)
$$X(t) = \int^t \sum_{i=1}^N f_i(t) g_i(u) dB(u)$$

with a proper canonical kernel $\sum_{i=1}^N f_i(t) g_i(u)$, where $\{f_i(t)\}$, $i=1, 2, \cdots$, N, satisfy

(II. 9) $\qquad \det(f_i(t_j)) \neq 0, \qquad$ *for any N different t_j,*

and $\{g_i(u)\}$, $i=1, 2, \cdots$, N, are linearly independent as the elements of $L^2(v \; ; \; t)^{2)}$ for every t.

Further the covariance function Γ of $Y(t)$ can be written in the form

$$\Gamma(s, t) = \sum_{i=1}^N f_i(t) h_i(s), \qquad s < t,$$

where $\{f_i(t)\}$, $i=1, 2, \cdots$, N, are the same as above and $\{h_i(s)\}$, $i=1, 2, \cdots$, N, are linearly independent.

Proof. By the assumptions there exists a proper canonical representation $(dB(t), F(t, u))$ of $Y(t)$:

$$Y(t) \sim X(t) = \int^t F(t, u) dB(u), \qquad (\sim : \text{equivalent in law})$$

It is sufficient to determine the form of $F(t, u)$ in the region $D = \{(u \; t,) \; ; \; u < t\}$ instead of $\bar{D}$ on account of the assumption ($\mathfrak{M}$. 4).

If $Y(t)$ is an N-ple Markov process, then we can prove that, for any $\{t_j\}$ with $t_1 < t_2 < \cdots < t_N$ and for any $\tau > t_N$ there exist $\{a_j(\tau \; ; \; t_1, t_2, \cdots, t_N)\}$ $j=1, 2, \cdots$, N, such that

(II. 10)
$$Y(\tau) - \sum_{j=1}^N a_j(\tau \; ; \; t_1, t_2, \cdots, t_N) Y(t_j)$$

is independent of every $Y(\sigma)$, $\sigma \leq t_1$. Therefore we have

---

2) $L^2(v \; ; \; t) = \{\varphi \; ; \; \varphi \in L^2(v) \text{ and } \varphi(u) \equiv 0 \text{ for } u > t\}$.

$$\int^{\sigma} F(\sigma, u)\{F(\tau, u) - \sum_{j=1}^{N} a_j(\tau\,;\, t_1, t_2, \cdots, t_N)F(t_j, u)\}\,dv(u) = 0\,,$$

using the representation of $Y(t)$. Since $F(\sigma, u)$ is a proper canonical kernel, it is equivalent to

(II. 11) $\qquad F(\tau, u) = \sum_{j=1}^{N} a_j(\tau\,;\, t_1, t_2, \cdots, t_N)F(t_j, u)$

as an element of $L^2(v\,;\, t_1)$ (see Theorem I. 7).

Take N different $\{s_j\}$ with $s_1 < s_2 < \cdots s_N$, arbitrarily in the interval $(-\infty,\ t_1) \cap T$. Expressing $F(\tau, u)$ and $\{F(t_j, u)\}$, $j=1, 2, \cdots, N$, in (II. 11) by $\{F(s_j, u)\}$, $j=1, 2, \cdots, N$, in the same way as in (II. 11), we get

$$\sum_{j=1}^{N} a_j(\tau\,;\, s_1, s_2, \cdots, s_N)F(s_j, u)$$

$$= \sum_{k,j=1}^{N} a_k(\tau\,;\, t_1, t_2, \cdots, t_N)a_j(t_k\,;\, s_1, s_2, \cdots, s_N)F(s_j, u)$$

as an element of $L^1(v\,;\, s_1)$.

Here $\{F(s_j, u)\}$, $j=1, 2, \cdots, N$, must be linearly independent functions in $L^2(v\,;\, s_1)$; in fact, if this is not true, then $\{E(Y(s_j)/\boldsymbol{B}_{s_1})\}$, $j=1, 2, \cdots, N$, are linearly dependent, which contradicts our assumption. Hence we have

(II. 12) $\qquad \sum_{k=1}^{N} a_k(\tau\,;\, t_1, t_2, \cdots, t_N)a_j(t_k\,;\, s_1, s_2, \cdots, s_N)$

$$= a_j(\tau\,;\, s_1, s_2, \cdots, s_N), \qquad \text{for every } j\,.$$

Now we can prove

(II. 13) $\qquad \det\,(a_j(t_k\,;\, s_1, s_2, \cdots, s_N)) \neq 0\,,$

because

$$F(t_j, u) = \sum_{k=1}^{N} a_k(t_j\,;\, s_1, s_2, \cdots, s_N)F(s_k, u), \qquad j = 1, 2, \cdots, N\,,$$

are linearly independent functions in $L^2(v\,;\, s_1)$. Therefore we have

(II. 14) $\qquad\qquad \boldsymbol{a}(\tau, \boldsymbol{t}) = \boldsymbol{a}(\tau, \boldsymbol{s})B(\boldsymbol{s}, \boldsymbol{t})$

by (II. 12) and (II. 13), where

$$\boldsymbol{a}(\tau, \boldsymbol{s}) = (a_1(\tau\,;\, s_1, s_2, \cdots, s_N), \cdots, a_N(\tau\,;\, s_1, s_2, \cdots, s_N))$$

and

$$B(\boldsymbol{t}, \boldsymbol{s}) = (b_{jk}(t_1, t_2, \cdots, t_N\,;\, s_1, s_2, \cdots, s_N)), \qquad j, k = 1, 2, \cdots, N,$$

with det $(B(\boldsymbol{t}, \boldsymbol{s})) \neq 0$. Taking N different $s_i'(< s_1)$, we have

$$\boldsymbol{a}(\tau, \boldsymbol{t}) = a(\tau, \boldsymbol{s})B(\boldsymbol{s}, \boldsymbol{t}) = \boldsymbol{a}(\tau, \boldsymbol{s}')B(\boldsymbol{s}', \boldsymbol{s})B(\boldsymbol{s}, \boldsymbol{t}),$$
$$\boldsymbol{a}(\tau, \boldsymbol{t}) = a(\tau, \boldsymbol{s}')B(\boldsymbol{s}', \boldsymbol{t}),$$

by (II. 14). Hence we have

(II. 15) $B(\boldsymbol{s}', \boldsymbol{s})B(\boldsymbol{s}, \boldsymbol{t}) = B(\boldsymbol{s}', \boldsymbol{t})$.

Fix all t_j's and define $\boldsymbol{f}_s(\tau)$, $\boldsymbol{s} = (s_1, s_2, \cdots, s_N)$, by

$$\boldsymbol{f}_s(\tau) = \boldsymbol{a}(\tau, \boldsymbol{s})B(\boldsymbol{s}, \boldsymbol{t}) \qquad \text{for} \quad \tau > s_N,$$

where $\boldsymbol{s}$ is any N-ple $(s_1, s_2, \cdots, s_N)$ such that $t_N > t_{N-1} \cdots > t_1 > s_N > s_{N-1} > \cdots > s_1$. Then we can use (II. 15) to see that $\boldsymbol{f}_{s'}$ is an extension of $\boldsymbol{f}_s(\tau)$ if $s_N > s_{N-1} > \cdots > s_1 > s_N' > s_{N-1}' > \cdots > s_1'$. Hence there exists a common extension for all $\boldsymbol{f}_s(\tau)$'s. We denote this common extension with $\boldsymbol{f}(t) = (f_1(t), \cdots, f_N(t))$. Obviously these $f_i(t)$ satisfy (II. 9) on account of (II. 13) and the definition of $\boldsymbol{f}_s(\tau)$.

Take $u \in T^0$ and fix it. If $\tau > t_N > \cdots > t_1 > s_N > \cdots > s_1 > u$, then we have

$$F(\tau, u) \Big(\equiv \sum_{j=1}^{N} a_j(\tau ; t_1, t_2, \cdots, t_N)F(t_j, u) \Big)$$
$$= \boldsymbol{a}(\tau, \boldsymbol{t})\boldsymbol{F}(\boldsymbol{t}, u)^* \qquad (\boldsymbol{F}(\boldsymbol{t}, u) = (F(t_1, u), \cdots, F(t_N, u))$$
$$= \boldsymbol{f}(\tau)B(\boldsymbol{s}, \boldsymbol{t})^{-1}\boldsymbol{F}(\boldsymbol{t}, u)^*$$
$$= \boldsymbol{f}(\tau)\boldsymbol{g}(u, \boldsymbol{s}, \boldsymbol{t})^*, \qquad (\boldsymbol{g}(u, \boldsymbol{s}, \boldsymbol{t}) = \boldsymbol{F}(\boldsymbol{t}, u)B(\boldsymbol{s}, \boldsymbol{t})^{*-1}).$$

For $\tau > t_N' > \cdots > t_1' > s_N' \cdots > s_1'$, this is equal to

$$\boldsymbol{f}(\tau)\boldsymbol{g}(u, \boldsymbol{s}', \boldsymbol{t}')^*,$$

so that

$$\boldsymbol{f}(\tau)\boldsymbol{g}(u, \boldsymbol{s}, \boldsymbol{t})^* = \boldsymbol{f}(\tau)\boldsymbol{g}(u, \boldsymbol{s}', \boldsymbol{t}')^*$$

for $\tau > \max(t_N', t_N)$. Since $\boldsymbol{f}$ satisfies (II. 9), we have

$$\boldsymbol{g}(u, \boldsymbol{s}, \boldsymbol{t}) = \boldsymbol{g}(u, \boldsymbol{s}', \boldsymbol{t}').$$

Therefore $\boldsymbol{g}(u) = \boldsymbol{g}(u, \boldsymbol{s}, \boldsymbol{t})$ is well defined as a function of u, and

$$F(t, u) = \boldsymbol{f}(t)\boldsymbol{g}(u)^* = \sum_{i=1}^{N} f_i(t)g_i(u),$$

where $\{g_i(u)\}$, $i = 1, 2, \cdots, N$, are linearly independent as elements of $L^2(v ; t)$, since $\{F(t_j, u)\}$, $j = 1, 2, \cdots, N$, are linearly independent.

Further we have

$$\Gamma(t, s) = \sum_{i=1}^{N} f_i(t)(\sum_{j=1}^{N} f_j(s) \int^s g_i(u)g_j(u)dv(u))$$

$$\equiv \sum_{i=1}^{N} f_i(t)h_i(s) \, .$$

Then if

$$\sum_{i=1}^{N} a_i h_i(s) \equiv 0$$

for some constants $a_1, a_2, \cdots, a_N$,

$$\int^s (\sum_{j=1}^{N} f_j(s)g_j(u))(\sum_{i=1}^{N} a_i g_i(u))dv(u) \equiv 0 \, .$$

Noting that $\sum_{j=1}^{N} f_j(s)g_j(u)$ is a proper canonical kernel, we have

$$\sum_{i=1}^{N} a_i g_i(u) = 0 \, .$$

Hence all the a_i must be 0. Thus we have proved the theorem completely.

A kernel $\sum_{i=1}^{N} f_i(t)g_i(u)$ satisfying the conditions stated in Theorem II. 2 is called a *Goursat kernel* of order N.

It should be noted that the expression (II. 13) is not uniquely determined, but the number of the summand is independent of the special way of expression as we have seen in the proof above.

As another remark, we should note that a process with a version of the form (II. 9) is not always an N-ple Markov process. In order that the converse of this theorem holds, it is sufficient to impose some regularity condition on the kernel as we shall see in § II. 4. 3°).

Example II. 1. If $f(t)$ is a function which is 1 for rational t and 0 for irrational t, then $X(t) = f(t)B_0(t)$, $0 \le t < \infty$, is not a 1-ple Markov process, though it is expressed in the form

$$X(t) = \int_0^t f(t)dB_0(u) \, .$$

§ II. 3. **Stationary multiple Markov Gaussian processes.**

Let $Y(t)$, $t \in T = (-\infty, \infty)$, be a stationary Gaussian process with mean 0 satisfying the conditions ($\mathfrak{M}. 1$), ($\mathfrak{M}. 2$) and ($\mathfrak{M}. 4'$).

($\mathfrak{M}. 4'$) $Y(t)$ is continuous in the mean.

Then by Karhunen's theory, we can see that $Y(t)$ has a canonical representation and that it is expressed in the form

$$Y(t) = \int_{-\infty}^{t} F(t-u)dB_0(u), \qquad E(dB_0(u))^2 = du,$$

using a canonical kernel $F(t-u)$. This canonical kernel is uniquely determined up to sign and is proper canonical.[3] Thus, in order to study the multiple Markov process for stationary case, it is sufficient for us to study the canonical kernel $F(t-u)$.

Lemma II. 1. *Let* $\{f_i(t)\}$, $i = 1, 2, \cdots, N$, *satisfy* (II. 9) *and let* $\{g_i(u)\}$, $i = 1, 2, \cdots, N$, *be linearly independent as elements of* $L^2((-\infty, c])$ *for every* c. *If* $\sum_{i=1}^{N} f_i(t)g_i(u)$ *is a function of* $t-u$ *in the domain* $\bar{D} = \{(u, t) ; u \leq t\}$, *then* $\{f_i(t)\}$ *is a fundamental system of solutions of a certain linear differential equation of order* N *with constant coefficients, and* $\{g_i(u)\}$ *is also a fundamental system of solutions of its adjoint differential equation.*

Proof. First we consider $F(t-u) = \sum_{i=1}^{N} f_i(t)g_i(u)$ in the region $D_0 = \{(u, t) ; u \leq 0, t \geq 0\}$. Let $\mathfrak{D}_0$ be the set of all C^∞-functions whose carriers are compact sets lying in the interval $(-\infty, 0]$. Then

$$(F*\varphi)(t) = \int_{-\infty}^{\infty} F(t-u)\varphi(u)du$$

is well defined by the assumption and belongs to $C^\infty((0, \infty))$ for every $\varphi \in \mathfrak{D}_0$ (Schwartz [1]).

Next we shall prove that there exist functions $\varphi_j(u)$, $j = 1, 2, \cdots, N$ in $\mathfrak{D}_0$ such that

(II. 16) $\det((g_i, \varphi_j)) \neq 0, \qquad i, j = 1, 2, \cdots, N,$

where (g, φ) denotes the inner product of g and φ in $L^2(T)$. In fact there exists a function $\varphi_1(u) \in \mathfrak{D}_0$ such that $(g_1, \varphi_1) \neq 0$. (If there were no such function, $g_1(u)$ must vanish on $(-\infty, 0)$). Inductively, suppose that $\varphi_1, \varphi_2, \cdots, \varphi_n \in \mathfrak{D}_0$ are chosen so that

$$\det((g_i, \varphi_j)) \neq 0, \qquad i, j = 1, 2, \cdots, n.$$

And consider the determinant

---

3) For proof see Karhunen [1]. Also, M. Nisio gave another proof, which I knew by private communication.

$$\begin{vmatrix} (g_1, \varphi_1) & (g_1, \varphi_2) & \cdots & (g_1, \varphi_n) & (g_1, \varphi) \\ (g_2, \varphi_1) & (g_2, \varphi_2) & \cdots & (g_2, \varphi_n) & (g_2, \varphi) \\ \multicolumn{5}{c}{\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots} \\ (g_{n+1}, \varphi_1) & (g_{n+1}, \varphi_2) & \cdots & (g_{n+1}, \varphi_n) & (g_{n+1}, \varphi) \end{vmatrix}.$$

If this determinant vanishes for every $\varphi \in \mathfrak{D}_0$, we get

$$\Delta_1(g_1, \varphi) + \Delta_2(g_2, \varphi) + \cdots + \Delta_{n+1}(g_{n+1}, \varphi) = 0, \qquad \varphi \in \mathfrak{D}_0,$$

by expanding this determinant with respect to the last column; since φ is arbitrary, we have

$$\Delta_1 g_1(u) + \Delta_2 g_2(u) + \cdots + \Delta_{n+1} g_{n+1}(u) = 0 \qquad \text{a.e. in } (-\infty, 0).$$

This contradicts the assumption that $\{g_i(u)\}$, $i = 1, 2, \cdots, n+1$, are linearly independent, since $\Delta_{n+1} \neq 0$ by the assumption of induction. Thus we can take $\{\varphi_j(u)\}$, $j = 1, 2, \cdots, N$, so that (II. 16) holds.

On the other hand, considering

$$(F * \varphi_j)(t) = \sum_{i=1}^{N} (g_i, \varphi_j) f_i(t)$$

and (II. 16), we can see that $f_i(t)$ is a linear combination of $(F * \varphi_j)(t)$ which belongs to $\boldsymbol{C}^\infty((0, \infty))$ Hence $f_i(t) \in \boldsymbol{C}^\infty((0, \infty))$ for every i, so that $F(t) \in \boldsymbol{C}^\infty((0, \infty))$. From these facts we can see that $g_i(u) \in \boldsymbol{C}^\infty((-\infty, 0))$ for every i.

Applying similar arguments to every region $D_a = \{(u, t) ; u \leq a, t \geq a\}$, $a \in T$, we can see that

$$f_i(\cdot), \ g(\cdot) \in \boldsymbol{C}^\infty(T^\circ), \qquad i = 1, 2, \cdots, N.$$

Thus we have

(II. 17) $$\sum_{i=1}^{N} f_i^{(k)}(t) g_i(u) = \frac{\partial^k}{\partial t^k} F(t - u) = (-1)^k \frac{\partial^k}{\partial u^k} F(t - u).$$

$$= (-1)^k \sum_{i=1}^{N} f_i(t) g_i^{(k)}(u), \qquad k = 0, 1, \cdots, N.$$

Putting $u = 0$, we get

$$\frac{d^k}{dt^k} F(t) = (-1)^k \sum_{i=1}^{N} f_i(t) g_i^{(k)}(0), \qquad k = 0, 1, \cdots, N.$$

Therefore there exist $b_0, b_1, \cdots, b_N$ such the $\sum_{i=1}^{N} |b_i| > 0$ and that

$$b_0 F^{(N)}(t) + b_1 F^{(N-1)}(t) + \cdots + b_N F(t) = 0, \qquad t > 0,$$

so that

$$b_0 F^{(N)}(t-u) + b_1 F^{(N-1)}(t-u) + \cdots + b_N F(t-u) = 0, \qquad t > u ,$$

namely

$$\sum_{i=1}^{N} (b_0 f_i^{(N)}(t) + b_1 f_i^{(N-1)}(t) + \cdots + b_N f_i(t)) g_i(u) = 0, \qquad t > u .$$

Since $\{g_i(u)\}$ are linearly independent in $L^2((-\infty, t])$,

$$b_0 f_i^{(N)}(t) + b_1 f_i^{(N-1)}(t) + \cdots + b_N f_i(t) = 0, \qquad i = 1, 2, \cdots, N .$$

If $b_0 = 0$, then $f_i(t)$ satisfies

$$c_1 f_1(t) + c_2 f_2(t) + \cdots + c_N f_N(t) = 0. \qquad \sum_{i=1}^{N} |c_i| \neq 0 ,$$

as a system of N solutions of linear differential equation of at most order $N-1$, which contradicts (II. 9). Hence $\{f_i(t)\}$ is a fundamental system of solutions of linear differential equation of order N with constant coefficients.

Exactly in the same way, we can prove the assertion for $\{g_i(u)\}$.

By the well-known fact in the theory of linear ordinary differential equations, $F(t-u)$ is a linear combination of the functions of the following types:

$$e^{-\lambda(t-u)} \sin \mu(t-u), \quad t^k u^{n-k} e^{-\lambda(t-u)} \sin \mu(t-u), \quad (\mu \neq 0)$$

(II. 18) $\quad e^{-\lambda(t-u)} \cos \mu(t-u), \quad t^k u^{n-k} e^{-\lambda(t-u)} \cos \mu(t-u), \quad (\mu \text{ may be } 0)$

$$0 \leq k \leq n, \quad n \leq N .$$

By Theorem II. 2 and Lemma II. 1 we have

Theorem II. 3. *If* $Y(t)$ *is a stationary N-ple Markov process satisfying the conditions* (𝔐. 1), (𝔐. 2) *and* (𝔐. 4′), *then its canonical kernel is a linear combination of the functions described in* (II. 18) *with* $\lambda > 0$.

The functions in (II. 18) (for $\mu \neq 0$) are split into two terms of the form $f(t)g(u)$. Therefore the number of the terms in the expression of the kernel is exactly N.

Corollary. *The spectral measure of a stationary N-ple Markov process is absolutely continuous with a density function of the following type*:

$$|Q(i\lambda)/P(i\lambda)|^2 ,$$

where P *is a polynomial of degree* N *and* Q *is also a polynomial of degree at most* $N-1$.

Proof. The spectral density function is obtained by Fourier transform of $F(\cdot)$. Hence our assertion is obvious.

This process is a component process of an N-dimensional stationary simple markov process in Doob's sense, Doob [1].

§II. 4. Some special multiple Markov Gaussian processes.

1°) Let $Y(t)$ be a stationary N-ple Markov Gaussian process which is differentiable (with respect to $L^2(\Omega)$-norm) up to $N-1$ times. Such process plays an important role in the study of N-ple Markov processes as is seen in Doob's work [1].

Now let us assume that $Y(t)$ is expressed in the form

$$Y(t) \sim X(t) = \int_{-\infty}^{t} F(t-u)dB_0(u)$$

with a proper canonical kernel

$$F(t-u) = \sum_{i=1}^{N} f_i(t)g_i(u) .$$

Then we have the following

Theorem II. 4. *Let $X(t)$ be a stationary N-ple Markov process. Then*

i) *a necessary and sufficient condition that $X(t)$ is differentiable is $F(0)=0$,*

ii) *in this case, there exists a complex number λ such that*

(II. 19) $$e^{\lambda t} \frac{d}{dt} e^{-\lambda t} X(t)$$

exists and it is a stationary $(N-1)$-ple Markov process.

Proof. i) If $h>0$,

$$\frac{1}{h}(X(t+h) - X(t)) = \frac{1}{h} \int_{t}^{t+h} F(t+h-u)dB_0(u)$$
$$+ \frac{1}{h} \int_{-\infty}^{t} \{F(t+h-u) - F(t-u)\} dB_0(u) .$$

Since $F(t-u)$ is analytic in D, the first term of the right hand side tends to 0 (in the mean) as h tends to 0 under the assumption $F(0)=0$. Hence $\lim_{h \to 0+} h^{-1}(X(t+h) - X(t))$ exists. Similarly $\lim_{h \to 0-} h^{-1}(X(t+h) - X(t))$ exists and

(II. 20) $$X'(t) = \int_{-\infty}^{t} \frac{\partial}{\partial t} F(t-u)dB_0(u) .$$

Conversely, if $X(t)$ is differentiable, then

$$dX(t) = F(0)dB_0(t) + dt \int_{-\infty}^{t} \frac{\partial}{\partial t} F(t-u)dB_0(u)$$

will be of order dt, so that $F(0)$ should vanish.

ii) As we have seen in Theorem II.3, $f_i(t)$ is a solution of a linear differential equation with constant coefficients. If we choose one of the characteristic roots of the differential equation, say λ,

$$e^{\lambda t} \frac{d}{dt} e^{-\lambda t} F(t-u)$$

is obviously a proper canonical Goursat kernel of order $N-1$. The existence of (II.19) and the stationary property are obvious.

When λ is real (II.19) is real valued process. When λ is complex, say $\lambda = \lambda_1 + i\lambda_2$, (II.19) is complex valued process, but

$$f_\lambda(t) \frac{d}{dt} f_\lambda(t)^{-1} X(t), \qquad f_\lambda(t) = e^{\lambda_1 t} \cos \lambda_2 t \ ,$$

is a real valued stationary process.

If $F(t)$ satisfies the conditions

$$F(0) = F'(0) = \cdots = F^{(N-1)}(0) = 0 \ ,$$

$X(t)$ is differentiable $N-1$ times. Then we can take a sequence of complex numbers $\lambda_1, \lambda_2, \cdots, \lambda_{N-1}$ such that

(II. 21) $$e^{\lambda_i t} \frac{d}{dt} e^{(\lambda_{i-1} - \lambda_i)t} \cdots \frac{d}{dt} e^{(\lambda_1 - \lambda_2)t} \frac{d}{dt} e^{-\lambda_1 t} X(t) \equiv X^{[i]}(t)$$

exists and it is a stationary $(N-i)$-ple Markov process.

Such a process was studied by Doob [1] and the formula (II.21) suggests more general differential operator which will appear in 3°)

2°) We shall now discuss a multiple Markov process with a homogeneous canonical kernel; $F(t, u)$ is called to be homogeneous function of degree α if $F(ct, cu) = c^\alpha F(t, u)$. This process can be transformed into a stationary process by time change by virtue of the following

Lemma II. 2.   (*P. Lévy*) *Let $X(t)$ be expressed as*

(II. 22) $$X(t) = \int_0^t F(t, u)dB_0(u)$$

with a proper canonical homogeneous (of degree α) kernel $F(t, u)$.

Then $t^{-\alpha-1/2}X(t)$ *is a stationary process of log t* (*P. Lévy* [4 ; *p.* 141]).

Applying this lemma to N-ple Markov process, we get

Theorem II. 5. *Let* $X(t)$ *be an* N-*ple Markov process and be expressed as in Lemma* II. 2. *Then* $e^{-(2\alpha+1)t}X(e^{2t})$ *is a stationary* N-*ple Markov process. Conversely any stationary* N-*ple Markov process* $\tilde{X}(t)$ *is an* N-*ple Markov process with homogeneous kernel of degree* 0 *changing the time parameter from t to* e^t :

$$\sqrt{t}\,\tilde{X}((\log t)/2) = X(t)\,.$$

Proof. The first part of the theorem is an immediate consequence of Lévy's lemma, if we notice that N ple Markov property is invariant under such time change.

If $X(t)$ is a stationary N-ple Markov process, it is the sum of the processes of the following types :

(II. 23) $$\int_{-\infty}^{t} e^{-\lambda(t-u)}dB_0(u), \quad \int_{-\infty}^{t} (t-u)^k e^{-\lambda(t-u)}dB_0(u)\,.$$

Changing the time scale, they become

(II. 23') $$\frac{1}{\sqrt{2}}\int_{0}^{t}(u/t)^{(\lambda-1)/2}d\tilde{B}_0(u), \quad \left(\frac{1}{\sqrt{2}}\right)^{k+1}\int_{0}^{t}(\log{(u/t)})^k(u/t)^{(\lambda-1)/2}d\tilde{B}_0(u)$$

respectively. Hence $X(t)$ is an N-ple Markov process with homogeneous kernel of degree 0. If λ is complex, these expressions are not real, but we can reduce them to real ones by the same procedure as used in the proof of Theorem II. 4. i).

3°) Let us generalize the results obtained in 1°) and 2°) to the case in which $X(t)$ is a general N-ple Markov process : our results include also those which were discussed by Dolph-Woodbury [1] and Lévy [4]. For the sake of brevity we shall assume $T=[0, \infty)$. We can discuss stationary N-ple Markov processes with parameter $\in(-\infty, \infty)$ in this scheme, if we apply to it the time change used in 2°).

We shall consider a process $X(t)$ which is expressed in the form

(II. 24) $$X(t) = \int_{0}^{t}\sum_{i=1}^{N} f_i(t)g_i(u)dB_0(u)\,,$$

with a proper canonical Goursat kernel $\sum_{i=1}^{N} f_i(t)g_i(u)$.

Hereafter (throughout this article), we shall always impose the following conditions on the kernel :

(A. 1) $f_i,\ g_i \in \boldsymbol{C}^\infty(T^\circ),$ for every i,

(A. 2) f_i and $W(g_1, g_2, \cdots, g_i)$ never vanish for every i,

where $W(g_1, g_2, \cdots, g_i)$ is the Wronskian of $\{g_j\}$, $j=1, 2, \cdots, i$.

By these assumptions, we can find functions $v_0, v_1, \cdots, v_{N-1}$ such that

(II. 25) $g_i(u) = (-1)^{N-i} v_0(u) \displaystyle\int_0^u v_1(u_1) \int_0^{u_1} v_2(u_2) \int_0^{u_2} \cdots$

$$\int_0^{u_{N-i-1}} v_{N-i}(u_{N-i})(du)^{N-i}$$

and that

(II. 26) $\begin{cases} v_i(u) \in \boldsymbol{C}^\infty(T^\circ), \\ v_i(u) \text{ never vanishes}, \quad i = 0, 1, \cdots, N-1. \end{cases}$

Using these functions $\{v_i(u)\}$, $i = 0, 1, \cdots, N-1$, and a function $v_N(u)$ satisfying (II. 26), we can define measures

$$m_i(M) = \int_M v_i(u)du, \quad M \in \boldsymbol{B}_T, \quad i = 0, 1, \cdots, N,$$

and the following differential operators:

$$L_t = \frac{d}{dm_0}\frac{d}{dm_1} \cdots \frac{d}{dm_{N-1}} \cdot \frac{1}{v_N(t)} \cdot,$$

$$L_t^{(j)} = \frac{d}{dm_j}\frac{d}{dm_{j+1}} \cdots \frac{d}{dm_{N-1}} \cdot \frac{1}{v_N(t)} \cdot,$$

$$L_u^* = \text{the adjoint operator of } L_t$$

$$\equiv \frac{d}{dm_N}\frac{d}{dm_{N-1}} \cdots \frac{d}{dm_1} \cdot \frac{1}{v_0(u)} \cdot,$$

$$L_u^{*(j)} = \frac{d}{dm_{N-j}}\frac{d}{dm_{N-j-1}} \cdots \frac{d}{dm_1} \cdot \frac{1}{v_0(u)} \cdot.$$

We shall often use the following notations:

$$F(t, u) = \sum_{i=1}^N f_i(t)g_i(u), \quad u \leqq t,$$

$$F^{(i)}(t, u) = \frac{\partial^i}{\partial t^i} F(t, u),$$

$$F^{[j]}(t, u) = L_t^{(N-j)} F(t, u).$$

Theorem II. 6. *Let $X(t)$ be defined by* (II. 24) *and let the canonical kernel of its representation satisfies the conditions* (A.1) *and* (A.2). *If*

(II. 27) $F(t, t) \equiv F^{(1)}(t, t) \equiv \cdots F^{(N-2)}(t, t) \equiv 0$ *and* $F^{(N-1)}(t, t)$
never vanishes,

then we have

i) $X^{(i)}(t) = \dfrac{d^i}{dt^i} X(t)$ *exists for every* $i \leqq N-1$,

ii) $X(t)$ *satisfies the equation*

(II. 28) $L_t X(t) = B_0'(t)$,

where $B_0'(t)$ *is a derivative of* $B_0(t)$ *in the symbolic sense, so that* (II. 28)
means $dL_t^{(1)} X(t) = v_0(t) dB_0(t)$, *and the measure* m_N *associated witn* L_t
should be taken appropriately.

   Proof. i) is proved in the same way as in the stationary case
(Theorem II. 4, i)).
   ii) Define $v_N(t) = f_1(t)$. Since $v_N(t)^{-1} \cdot F(t,t) \equiv 0$ by assumption,
we can prove the existence of $\dfrac{d}{dt} v_N(t)^{-1} X(t)$; namely $L_t^{(N-1)} X(t)$ ex-
ists. Similarly, we can prove the existence of $L_t^{(i)} X(t)$, $i = N-2$,
$N-3, \cdots$, 1, since $F^{[1]}(t, t) \equiv F^{[2]}(t, t) \equiv \cdots \equiv F^{[N-2]}(t, t) \equiv 0$.
   Rewriting (II. 27) in the following forms

$$\sum_{i=1}^{N} f_i^{(k)}(t) g_i(t) = 0, \qquad k = 0, 1, \cdots, N-1,$$

$$\sum_{i=1}^{N} f_i^{(N-1)}(t) g_i(t) = a(t), \qquad \text{with } a(t) \equiv F^{(N-1)}(t, t),$$

we can see that $F(t, u)/a(u)$ is a Riemann function for a certain
linear differential equation

$$\widetilde{L}_t f = 0$$

of order N. The fundamental system of solutions of its adjoint
differential equation

$$\widetilde{L}_u^* g = 0$$

is $\{g_i(u)/a(u)\}$, $i = 1, 2, \cdots, N$, as is well known. Hence $\widetilde{L}_u^* = L_u^* \cdot v(u)$
with a certain function $v(u)$. Thus we can prove that $\widetilde{L}_t = v(t) \cdot L_t$.
By the property of Riemann function $v(t)$ must be 1, and therefore
we have

(II. 29) $f_i(t) = v_N(t) \displaystyle\int_0^t \mathrm{dm}_{N-1} \int \mathrm{dm}_{N-2} \int \cdots \int \mathrm{dm}_{N-i+1}$,

$$i = 1, 2, \cdots, N,$$

which proves that

$$L_t^{(1)}F(t, u) = g_N(u) \equiv v_0(u) .$$

Hence

$$L_t^{(1)}X(t) = \int_0^t L_t^{(1)}F(t, u)dB_0(u) = \int_0^t v_0(u)dB_0(u) .$$

Thus we can prove ii).

Combining this Theorem II. 6 with Theorem II. 9. in § II. 5 we can see that this $X(t)$ process is an N-ple *Markov process in the restricted sense* in Lévy's terminology

Corollary. *Under the same assumptions as in Theorem* II. 6, $L_t^{(i)}X(t)$ *is an* $(N-i)$-*ple Markov process in the restricted sense.*

Proof. As is easily seen in the proof of the theorem above,

$$L_t^{(i)}X(t) = \int_0^t L_t^{(i)}F(t, u)dB_0(u) .$$

Since $L_t^{(i)}F(t, u)$ is a Riemann function for the differential equation

$$\frac{d}{dm_1} \cdots \frac{d}{dm_{i-1}} f = 0 ,$$

our assertion is obvious.

Theorem II. 7. *Let* $\{v_i(u)\}$, $i = 0, 1, \cdots, N$, *be functions satisfying the condition* (II. 26). *If we define* $f_i(t)$'*s and* $g_i(u)$'*s by* (II. 29) *and* (II. 25) *respectively, then*

i) $F(t, u) = \sum_{i=1}^{N} f_i(t)g_i(u)$ *is a proper canonical kernel,*

ii) *a process defined by*

$$X(t) = \int_0^t \sum_{i=1}^{N} f_i(t)g_i(u)dB_0(u)$$

is an N-*ple Markov process,*

iii) $F(t, t) \equiv F^{(1)}(t, t) \equiv \cdots \equiv F^{(N-2)}(t, t) \equiv 0$ *and* $F^{(N-1)}(t, t)$ *never vanishes*; *namely Theorem* II. 6 *holds for this process.*

Proof. i). We shall prove i) by using the kernel criterion which was given in § I. 4.

Suppose that

$$\int_0^t F(t, u)\varphi(u)du \equiv 0 \qquad \text{in } (0, t_0)$$

for some $t_0 \in T$ and some $\varphi \in L^2([0, t_0])$. Writing it in the form

$$\sum_{i=1}^{N} f_i(t)\int_0^t g_i(u)\varphi(u)du \equiv 0 ,$$

and multiplying $v_N(t)^{-1}$, we can obtain its derivative (in Radon-Nikodym sense)

$$\sum_{i=2}^{N} f_i^{[1]}(t)\int_0^t g_i(u)\varphi(u)du + (\sum_{i=1}^{N} f_i(t)g_i(t))\varphi(t) = 0$$

$$\text{a.e. in } (0, t_0).$$

Since the second term of the left hand side vanishes (because $\sum_{i=1}^{N} f_i(t)g_i(u)$ is a Riemann function corresponding to L_t), we have

$$\sum_{i=1}^{N} f_i^{(1)}(t)\int_0^t g_i(u)\varphi(u)du \equiv 0, \quad \text{in } (0, t_0),$$

by taking an appropriate version.

Repeating such procedures, we can prove

$$v_1(t)\int_0^t v_0(u)\varphi(u)du \equiv 0, \quad \text{in } (0, t_0).$$

Hence $\varphi(u)$ must be 0 as an element of $L^2([0, t_0])$.

ii) First we shall prove that

(II. 30) $\Delta(t_1, t_2, \cdots, t_N) \equiv \det (f_i(t_j)) \neq 0$ for any different t_j's.

If $\Delta(t_1, t_2, \cdots, t_N) = 0$ for some $t_1 < t_2 < \cdots < t_N$, we can prove

$D(t_1, t_2, \cdots, t_N)$

$$\equiv \begin{vmatrix} \int_{t_2}^{t_1} v_{N-1}du, & \int_{t_3}^{t_2} v_{N-1}du, & \cdots, \int_{t_N}^{t_{N-1}} v_{N-1}du \\ \int_{t_2}^{t_1} v_{N-1}\int v_{N-2}(du)^2, & \int_{t_3}^{t_2} v_{N-1}\int v_{N-2}(du)^2, \cdots, \int_{t_N}^{t_{N-1}} v_{N-1}\int v_{N-2}(du)^2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \int_{t_2}^{t_1} v_{N-1}\int \cdots \int v_1(du)^{N-1}, & \int_{t_3}^{t_2} v_{N-1}\int \cdots \int v_1(du)^{N-1}, \\ & \cdots, \int_{t_N}^{t_{N-1}} v_{N-1}\int \cdots \int v_1(du)^{N-1} \end{vmatrix} = 0,$$

since v_N never vanishes. By the mean value theorem, we have

$$\begin{vmatrix} v_{N-1}(t_1'), & v_{N-1}(t_2'), & \cdots, v_{N-1}(t_{N-1}') \\ v_{N-1}(t_1')\int_0^{t_1'}v_{N-2}du, & v_{N-1}(t_2')\int_0^{t_2'}v_{N-2}du, \cdots, v_{N-1}(t_{N-1}')\int_0^{t_{N-1}'}v_{N-2}du \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ v_{N-1}(t_1')\int_0^{t_1'}\cdots\int v_1(du)^{N-2}, & v_{N-1}(t_2')\int_0^{t_2'}\cdots\int v_1(du)^{N-2}, \\ & \cdots, v_{N-1}(t_{N-1}')\int_0^{t_{N-1}'}\cdots\int v_1(du)^{N-2} \end{vmatrix} = 0$$

for some $\{t'_i\}$, $i = 1, 2, \cdots, N-1$, with $t'_i \in (t_i, t_{i+1})$. Since v_{N-1} never vaishes, it is proved that

$$D(t'_1, t'_2, \cdots, t'_{N-1}) = 0 .$$

Successively we have

$$D(t''_1, t''_2, \cdots, t''_{N-2}) = \cdots = D(t_1^{(N-2)}, t_2^{(N-2)}) = 0$$

for some $\{t_i^{(k)}\}$'s with $t_i^{(k)} \in (t_i^{(k-1)}, t_{i+1}^{(k-1)})$.

Finally we have

$$\int_{t_2^{(N-2)}}^{t_1^{(N-2)}} v_1(u)du = 0 ,$$

which contradicts the assumption (II. 26).

Therefore, by (II. 30), we can find functions $\{a_j(t ; t_1, t_2, \cdots, t_N)\}$, $j = 1, 2, \cdots, N$, such that

(II. 31) $$\sum_{j=1}^{N} a_j(t ; t_1, t_2, \cdots, t_N) f_i(t_j) = f_i(t)$$

holds for every i. Hence

$$X(t) - \sum_{j=1}^{N} a_j(t ; t_1, t_2, \cdots, t_N) X(t_j)$$

is independent of every $X(\tau)$; $\tau \leq t_1$.

On the other hand $\{E(X(t_j)/\boldsymbol{B}_{t_0})\}$, $j = 1, 2, \cdots, N$, is linearly independent for any choice of t_j's with $t_0 \leq t_1 < \cdots < t_N$, since $g_i(u)$'s are linearly independent as elements of $\boldsymbol{L}^2([0, t_0])$. Thus the assertion ii) is proved.

iii) is easily proved, noting that $F(t, u)$ is the Riemann function corresponding to L_t.

Theorem II. 8. *Let $X(t)$ be defined by* (II. 24) *with a canonical kernel* $F(t, u) = \sum_{i=1}^{N} f_i(t) g_i(u)$, *and let $f_i(t)$'s and $g_i(u)$'s satisfy the conditions* (A. 1) *and* (A. 2). *If*

$$F(t, t) \equiv F^{(1)}(t, t) \equiv \cdots \equiv F^{(k-1)}(t, t) \equiv 0 \text{ and } F^{(k)}(t, t) \not\equiv 0$$

for some $k(< N-1)$ independent of t, then there exist $Y(t)$ which is an N-ple Markov process in the restricted sense and $(N-1)$-th order differential operator M_t such that

(II. 32) $$X(t) = M_t Y(t) .$$

Proof. By assumptions (A. 1) and (A. 2), $g_i(u)$'s can be expressed

in the form (II. 25). Therefore there exist a differential operator L_t and a Riemann function $R(t, u) = \sum_{i=1}^{N} \tilde{f}_i(t) g_i(u)$ corresponding to L_t, where $\tilde{f}_i(t)$ is defined by

$$\tilde{f}_i(t) = v_N(t) \int_0^t dm_{N-1} \int dm_{N-2} \int \cdots \int dm_{N-i+1}, \qquad i = 1, 2, \cdots, N.$$

Define $Y(t)$ by

$$Y(t) = \int_0^t R(t, u) dB_0(u).$$

Then, this $Y(t)$ will be the one to be obtained.

The assumption (A. 2) implies $W(\tilde{f}_1, \tilde{f}_2, \cdots, \tilde{f}_N) \neq 0$ for every t. Therefore we can find a differential operator M_t such that

$$M_t \tilde{f}_i(t) = \sum_{j=0}^{N-1} b_j(t) \tilde{f}_i^{(j)}(t) = f_i(t), \qquad i = 1, 2, \cdots, N.$$

Noting that $Y(t)$ has the j-th order derivative

$$Y^{(j)}(t) = \int_0^t R^{(j)}(t, u) dB_0(u)$$

for every $j \leq N - 1$, by Theorem II. 6, i), M_t can be operated to $Y(t)$ and we have

$$\begin{aligned}
M_t Y(t) &\equiv \sum_{j=1}^{N-1} b_j(t) Y^{(i)}(t) \\
&= \int_0^t \sum_{j=1}^{N-1} b_j(t) R^{(j)}(t, u) dB_0(u) \\
&= \int_0^t \sum_{i=1}^{N} (\sum_{j=1}^{N} b_j(t) \tilde{f}_i(t)) g_i(u) dB_0(u) \\
&= \int_0^t \sum_{i=1}^{N} f_i(t) g_i(u) dB_0(u)
\end{aligned}$$

This completes the proof.

Example II. 2. Lévy's example $X_1(t)$ which we discussed in Example 1.5 satisfies all assumptions imposed on the canonical kernel in Theorem II. 8, where $N = 2$, $v_0 = v_1 = 1$ and $k = 0$. In this case

$$Y(t) = \int_0^t (t - u) dB_1(u)$$

and $M_t = \dfrac{d}{dt} t \cdot.$

§II. 5. Prediction of multiple Markov Gaussian processes.

For a Gaussian process $Y(t)$, the least square linear prediction on the basis of its values before $s(<t)$ is obtained by the conditional expectation $E(Y(t)/\boldsymbol{B}_s)$, as is well known. If there exists a canonical representation $(dB(t), F(t, u))$ of $Y(t)$, then, by definition,

$$(\text{II. 33}) \qquad E(X(t)/\boldsymbol{B}_s) = \int^s F(t, u)dB(u), \quad (X(t) = \int^t F(t, u)dB(u)).$$

It is our aim to express it in terms of $X(\tau)$, $\tau \leq s$.

Theorem II. 9. *Let $Y(t)$ be a process defined in Theorem II. 8. Using the same notations, we have*

$$(\text{II. 34}) \qquad E(Y(t)/\boldsymbol{B}_s) = \sum_{j=1}^{N} b_j(t,s)Y^{[j-1]}(s), \qquad s < t,$$

where

$$(\text{II. 35}) \qquad\qquad b_j(t, s) = \sum_{j=1}^{N} \tilde{f}_i(t)\Delta_{ji}/\Delta(s)$$

with

$$\Delta(s) = \det(\tilde{f}^{[j-1]}_{(s)}(s)) \ \text{and} \ \Delta_{ji} = (j, i)\text{-cofactor of } \Delta(s).$$

Proof. Putting $U_i(s) = \int_0^s g_i(u)dB_0(u)$, we have

$$(\text{II. 36}) \qquad E(Y(t)/\boldsymbol{B}_s) = \int_0^s \sum_{i=1}^{N} \tilde{f}_i(t)g_i(u)dB_0(u) = \sum_{i=1}^{N} \tilde{f}_i(t)U_i(s).$$

On the other hand,

$$Y^{[k]}(s) = L_s^{(N-k)}Y(s) = \sum_{i=k+1}^{N} \tilde{f}^{[k]}_i(s)U_i(s), \qquad k = 0, 1, \cdots, N-1.$$

Since

$$\Delta(s) = (\tilde{f}^{[k]}_i(s))_{\substack{i=1,2,\cdots,N \\ k=0,1,\cdots,N-1}} = \prod_{i=1}^{N} v_i(s) \neq 0,$$

$U_i(s)$ can be written in terms of $Y^{[k]}(s)$'s, $k=0, 1, \cdots, N-1$, that is

$$(\text{II. 37}) \qquad U_i(s) = \frac{1}{\Delta(s)} \begin{vmatrix} \tilde{f}_1(s) & \tilde{f}_2(s) & \cdots & Y^i(s) & \cdots & \tilde{f}_N(s) \\ 0 & \tilde{f}^{[1]}_2(s) & \cdots & Y^{[1]}(s) & \cdots & \tilde{f}^{[1]}_N(s) \\ 0 & 0 & \cdots & Y^{[2]}(s) & \cdots & \tilde{f}^{[2]}_N(s) \\ \multicolumn{6}{c}{\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots} \\ 0 & 0 & \cdots & Y^{[N-1]}(s) & \cdots & \tilde{f}^{[N-1]}_N(s) \end{vmatrix}.$$

Combinig this with (II. 36), we have

$$E(Y(t)/\boldsymbol{B}_s) = \sum_{i=1}^{N} \tilde{f}_i(t) \left(\sum_{j=1}^{N} \frac{\Delta_{jt}}{\Delta(s)} Y^{[j-1]}(s) \right)$$

$$= \sum_{j=1}^{N} \left(\sum_{i=1}^{N} \tilde{f}_i(t) \frac{\Delta_{jt}}{\Delta(s)} \right) Y^{[j-1]}(s) ,$$

which was to be proved.

Corollary. *Let $X(t)$ and $Y(t)$ be the same processes as in Theorem II. 8. Then we have*

(II. 38) $$E(X(t)/\boldsymbol{B}_s) = \sum_{j=1}^{N} c_j(t, s) Y^{[j-1]}(s) ,$$

where

$$c_j(t, s) = \sum_{i=1}^{N} f_i(t) \frac{\Delta_{jt}}{\Delta(s)} ,$$

$\Delta(s)$ *and* Δ_{ji} *being the same as in* (II. 35).

Proof. Noting that

$$E(X(t)/\boldsymbol{B}_s) = \sum_{i=1}^{N} f_i(t) U_i(s) \qquad \text{for every } s < t ,$$

we can easily prove (II. 38).

This corollary suggests the following symbolic calculous of determing the predictor. Using the differential operator M_t defined in (II. 33), (II. 38) becomes

$$E(M_t Y(t)/\boldsymbol{B}_s) = \sum_{j=1}^{N} c_j(t, s) Y^{[j-1]}(s)$$

$$= \sum_{j=1}^{N} M_t b_j(t,s) Y^{[j-1]}(s) = M_t E(Y(t)/\boldsymbol{B}_s) .$$

This means that M_t and $E(\cdot/\boldsymbol{B}_s)$ are commutative.

On the other hand, (II. 38) may be denoted as

$$E(X(t)/\boldsymbol{B}_s) = \sum_{j=1}^{N} c_j(t, s) (M_t^{-1} X(t))_{t=s}^{[j-1]} ,$$

where M_t^{-1} is an integral operator such as $M_t(M_t^{-1} X(t)) = X(t)$. Hence formally speaking, the prediction operator for $X(t)$ is composed of differential and integral operators.

Theorem II. 10. *Under the same assumption as in Theorem II. 9,*

$$\lim_{s_N \uparrow s} \sum_{i=1}^{N} a_i(t, s_1, s_2, \cdots, s_N) Y(s_i)$$

exists and equals $E(Y(t)/B_s)$, where $a_i(t, s_1, s_2, \cdots, s_N)$, $i = 1, 2, \cdots, N$ are the functions determined by Theorem II. 7, (II. 31).

Proof. Refering to the proof of Theorem II. 7, we have

$$\sum_{i=1}^{N} a_i(t, s_1, s_2, \cdots, s_N) Y(s_i) = \sum_{i=1}^{N} \left(\frac{1}{\Delta_N(s_1, \cdots, s_N)} \sum_{j=1}^{N} \tilde{f}_j(t) \Delta_{ji}^{(N)} \right) Y(s_i)$$

$$= \sum_{j=1}^{N} \frac{1}{\Delta_N(s_1, \cdots, s_N)} \tilde{f}_j(t) (\sum_{i=1}^{N} \Delta_{ji}^{(N)} Y(s_i)),$$

where $\Delta_N(s_1, s_N, \cdots, s_N)$ has been defined in Lemma II. 5 and $\Delta_{ji}^{(N)}$ is its (j, i)-cofactor. Letting $s_1, s_2, \cdots, s_N$ tend to s successively, we can easily prove that $\sum_{i=1}^{N} a_i(t, s_1, s_2, \cdots, s_N) Y(s_i)$ tends to

$$\sum_{i=1}^{N} \frac{1}{\Delta(s)} \tilde{f}_i(t) \sum_{j=1}^{N} \Delta_{ji} Y^{[j-1]}(s) = \sum_{i=1}^{N} b_i(t, s) Y^{[i-1]}(s)$$

as was to be proved.

For stationry case, such prediction problem is well known (cf. J.L. Doob [1], [2])

§II. 6. Sum of stationary multiple Markov Gaussian processes.

As we discussed in §II. 3, any stationary N-ple Markov process is considered as the sum of stationary N_i-ple Markov processes in the restricted sense with $\sum N_i = N$. The converse problem will be discussed here.

For the sake of simplicity we shall consider the sum of stationary simple Markov processes, which is 1-ple Markov process in the restricted sense. General cases are treated similarly.

Let $Y_j(t)$, $j = 1, 2, \cdots, N$, be stationary simple Markov processes. Taking appropriate versions we can express $Y_j(t)$ with respect to the same random measure $B_0(\cdot)$ as follows

(II. 39) $\qquad Y_j(t) = c_j \int_{-\infty}^{t} e^{-\lambda_j(t-u)} dB_0(u), \qquad \lambda_j(>0), \ c_j : \text{constants}$

$$j = 1, 2, \cdots, N.$$

Now let us consider

(II. 40) $\qquad Y(t) = \sum_{j=1}^{N} Y_j(t) = \int_{-\infty}^{t} \sum_{j=1}^{N} c_j e^{-\lambda_j(t-u)} dB_0(u)$

Obviously it is at most N-ple Markov stationary process. Even in the cast that all the λ_j's are distinct, the kernel $F(t-u) = \sum_{j=1}^{N} c_j e^{-\lambda_j(t-u)}$

is not always canonical (Example II. 3), and $Y(t)$ is not always N-ple Markov process (Example III. 3.).

Let $\hat{F}(\lambda)$ be the Fourier transform of $F(x)$; $\hat{F}(\lambda) = \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-i\lambda x}$. $F(x)dx$. Then

(II. 41)
$$\hat{F}(\lambda) = \frac{Q(i\lambda)}{\prod\limits_{j=1}^{N}(i\lambda + \lambda_j)}, \quad (i = \sqrt{-1}),$$

where $Q(i\lambda)$ is the polynomial of $i\lambda$ at most of degree $N-1$. Writing

(II. 42)
$$Q(i\lambda) = \sum_{j=0}^{N-1} a_j (i\lambda)^{N-1-j},$$

we have

Theorem II. 11. *$F(t-u)$ is the proper canonical kernel if and only if $Q(x)$ has no zero point with positive real part.*

Proof. We use the kernel criterion proved in §I. 4. Define the numbers b_ν, $\nu = 0, 1, \cdots, N$, by

$$\prod_{j=1}^{N}(x+\lambda_j) = \sum_{\nu=0}^{N} b_\nu x^{N-\nu}.$$

And define the differential operator L_t by

$$L_t = \sum_{\nu=0}^{N} b_\nu \left(\frac{d}{dt}\right)^{N-\nu}.$$

Then we can easily see that

$$L_t e^{-\lambda_j(t-u)} = 0, \quad \text{consequently} \quad L_t F(u-t) = 0.$$

Now suppose

(II. 43) $$\int_{-\infty}^{t} F(t-u)\varphi(u) \equiv 0, \quad \text{for some} \quad \varphi \in L^2((-\infty, a]) \cap \mathfrak{D},$$

where $\mathfrak{D} = \{\varphi ; \varphi \in C^\infty \text{ and has compact carrier}\}$
The $(k+1)$-th derivative of it is

(II. 44) $$F(0)\varphi^{(k)}(t) + F'(0)\varphi^{(k-1)}(t) + \cdots + F^{(k)}(0)\varphi(t)$$
$$+ \int_{-\infty}^{t} F^{(k+1)}(t-u)\varphi(u)du = 0, \quad k = 0, 1, \cdots, N-1.$$

Then from (II. 43) and (II. 44), we have

$$F(0)\left(\sum_{k=0}^{N-1} b_{N-k-1}\varphi^{(k)}(t)\right) + F'(0)\left(\sum_{k=1}^{N-1} b_{N-k-1}\varphi^{(k-1)}(t)\right) + \cdots$$
$$+ F^{(N-1)}(0)b_0\varphi(t) + \int_{-\infty}^{t} L_t F(t-u)\varphi(u)du = 0,$$

that is,

$$\sum_{\nu=0}^{N-1} F^{(\nu)}(0)\left(\sum_{k=\nu}^{N-1} b_{N-k-1}\varphi^{(k-\nu)}(t)\right) = 0 .$$

If we introduce a new differential operator

$$\widetilde{L}_t = \sum_{j=0}^{N-1} \tilde{a}_j\left(\frac{d}{dt}\right)^{N-1-j} ,$$

with $\tilde{a}_j = \sum_{\nu=0}^{j} F^{(\nu)}(0)b_{j-\nu}$, then the above equality can be written as

(II. 45) $$\widetilde{L}_t\varphi(t) = 0 .$$

Non trivial function φ satisfying (II. 45) exists and belongs to $L^2((-\infty, a])$ if and only if the characteristic equation

(II. 46) $$\sum_{j=0}^{N-1} \tilde{a}_j x^{N-1-j} = 0$$

of $\widetilde{L}_t$ has at least one root with positive real part.

On the other hand, noting that

$$\int_0^\infty e^{-i\lambda x} F^{(k)}(x)dx = F^{(k-1)}(0) + (i\lambda)F^{(k-2)}(0) + (i\lambda)^2 F^{(k-3)}(0) + \cdots$$
$$+ (i\lambda)^{k-1}\hat{F}(\lambda)$$

and

$$\hat{F}(\lambda)\sum_{j=0}^{N} b_j(i\lambda)^{N-j} = Q(i\lambda) ,$$

we can prove

$$a_j = \tilde{a}_j \quad \text{for every } j ,$$

Hence the desired condition is equivalent to the one that

(II. 46′) $$\sum_{j=0}^{N-1} a_j x^{N-1-j} = Q(x) = 0 .$$

has no root with positive real part.

Generally, not assuming that $\varphi \in \mathfrak{D}$ in (II. 43), the same assertion is true, since $(F*\varphi)(t)=0$ is equivalent to $F*(\varphi*\alpha)(t)=0$ for every $\alpha \in \mathfrak{D}$. (Note that $(\varphi*\alpha)(t) \in C^\infty$). This completes the proof.

If we observe the proof of this theorem, we can see that its proposition may be improved to the case that all the $Y_j(t)$'s are stationary multiple Markov processes in the restricted sense.

As an obvious consequence of this theorem, we can say that $Y(t)$ defined by (II. 40) is an N-ple Markov stationary process, if the condition of the theorem is fullfilled. In particular, it is an

N-ple Markov process in the restricted sense if and only if all the a_j's are zero except a_0.

Example II. 3.　Consider a process

$$X(t) = 3\int_{-\infty}^{t} e^{-(t-u)}dB_0(u) - 4\int_{-\infty}^{t} e^{-2(t-u)}dB_0(u) .$$

The kernel $3e^{-(t-u)} - 4e^{-2(t-u)}$ is not a proper canonical kernel.

Sction III.　Lévy's M(t) process.

§ III. 1.　Definition and known results.

Let $X(A, \omega)$, $A \in \mathbf{E}^N$ (N-dimensional Euclidean space), $\omega \in \Omega$, be a *Brownian motion with a parameter space* $\mathbf{E}^N$, that is

(III. 1)
$$\begin{cases} \text{i)} & X(A) \text{ is a Gaussian random variable with mean } 0 \\ & \text{for every } A, \\ \text{ii)} & X(O) = 0, \text{ where } O \text{ is the origin of } \mathbf{E}^N, \\ \text{iii)} & E(X(A) - X(B))^2 = r(A, B), \text{ where } r(A, B) \text{ denotes the} \\ & \text{distance between } A \text{ and } B. \end{cases}$$

Since $X(A, \omega)$ is continuous in A for almost all ω, (P. Lévy [1], T. Sirao [1]), the following integral is well defined and we have a Gaussian process $M_N(t)$ with a parameter space $T = [0, \infty)$,

(III. 2)
$$M_N(t) = \int_{S(t)} X(A)d\sigma(A) ,$$

where $S_N(t)$ is the sphere in E^N with radius t and $d\sigma$ is the uniform measure on $S_N(t)$ with $\sigma(S_N(t)) = 1$.

P. Lévy studied the canonical representation and the Markov property of this process when N is odd (P. Lévy [3], [4]). Since $E(M_N(t)) = 0$, the covariance function of $M_N(t)$ is

(III. 3)
$$\begin{aligned} \Gamma_N(t, s) &= E(M_N(t)M_N(s)) \\ &= \int_{A \in S_N(t)} \int_{B \in S_N(s)} E\{X(A)X(B)\}d\sigma(A)d\sigma(B) \\ &= (t + s - \rho_N(t, s))/2 \end{aligned}$$

where

$$\rho_N(t, s) = \int_{A \in S_N(t)} \int_{B \in S_N(s)} r(A, B)d\sigma(A)d\sigma(B) .$$

By the simple computations we have, for $t = s$,

(III. 4)
$$\rho_N(t, t) = tJ_{N-2}/I_{N-2} \qquad (N \geqq 3)$$

with $I_k = \int_0^{\pi/2} \sin^k\theta \, d\theta$ and $J_k = \int_0^{\pi} \sin^k\theta \sin \frac{\theta}{2} d\theta$, and for $s \neq t$

(III.5) $\rho_N(t, s) = \frac{1}{2 I_{N-2}} \int_0^{\pi} r \sin^{N-2}\theta \, d\theta$ $(N \geq 3)$

with $r = (t^2 + s^2 - 2ts \cos \theta)^{1/2}$.

Using the analytic property of $\Gamma_N(t, s)$ and others, P. Lévy [4] obtained many important results concerning $M_N(t)$. First, if $N = 2p + 1$, $M_N(t)$ may be expressed as

(III.6) $M_N(t) = \int_0^t P_N(u/t) dB_0(u)$,

where $P_N\left(\frac{u}{t}\right)$ is a canonical kernal defined by

(III.7) $P_N(u) = \frac{2p}{\sqrt{\pi}} \sqrt{I_{2p}} \int_u^1 (1 - x^2)^{p-1} dx$

                          $=$ polynomial of degree $2p - 1$.

For example

Example III.1.

$$M_5(t) = \int_0^t (2/3 - u/t + u^3/3t^3) \sqrt{3} \, dB_0(u)$$

Example III.2.

$$M_7(t) = \int_0^t (2/5 - 3u/4t + u^3/2t^3 - 3u^5/20t^5) \sqrt{10} dB_0(u).$$

Concerning the Markov property, it was proved that $M_{2p+1}(t)$ has continuous derivatives of orders $1, 2, \cdots, p$ and it is a $(p+1)$-ple Markov process in the restricted sense.

§ III.2. Canonical representation of $M_N(t)$ process.

We are now interested in the canonical representation of $M_N(t)$ for the case that N is even particularly. First we shall consider some properties of $\Gamma_N(t, s)$ for odd and even N, and then we shall study $M_N(t)$ process.

As P. Lévy pointed out $e^{-t}M(e^{2t})$, which will be denoted by $X_N(t)$, becomes a stationary Gaussian process with parameter space $(-\infty, \infty)$. In fact

(III.8) $E(X_N(t)X_N(t+h)) = \frac{1}{2}\Big(2 \cosh h - \frac{1}{\sqrt{2} I_{N-2}} \int_0^{\pi} (\cosh (2h)$

                          $- \cos \theta)^{1/2} \sin^{N-2}\theta \, d\theta\Big)$ $h \geq 0$.

It is a function of h and will be denoted by $\gamma_N(h)$.

Lemma. *If $N \geq 4$, $\gamma_N(h)$ belongs to $\mathbf{C}^2$ and satisfies the following equation*

$$(\text{III. }9) \qquad (2N-3)^2\gamma_N(h) - \gamma_N''(h) = 4(N-1)(N-2)\gamma_{N-2}(h) \ .$$

Proof. If we note only the differentiability under the integral sign in the formula (III. 8), we can easily prove the existence of $\gamma_N'(h)$ and $\gamma_N''(h)$. Exact forms of them are

$$\gamma_N'(h) = \sinh h - \frac{\sinh 2h}{2I_{N-2}} \int_0^\pi \{2(\cosh 2h - \cos\theta)\}^{-1/2} \sin^{N-2}\theta \, d\theta$$

$$\gamma_N''(h) = \cosh h - \frac{1}{2\sqrt{2}\,I_{N-2}} \int_0^\pi \{2\cosh 2h(\cosh 2h - \cos\theta)$$

$$- \cosh^2 2h + 1\} \{\cosh 2h - \cos\theta\}^{-3/2} \sin^{N-2}\theta \, d\theta$$

Thus we obtain (III. 9).

Theorem III. 1. *If $N \geq 4$, we have*

$$(\text{III. }10) \qquad c_N X_{N-2}(t) = e^{-(2N-3)t} \frac{d}{dt} e^{(2N-3)t} X_N(t), \ c_N = 2\sqrt{(N-1)(N-2)} \ .$$

(Here a process and its version are identified)

Proof. From the above lemma, we can see that $e^{(2N-3)t}X_N(t)$ is differentiable. On the other hand, $X_N(t)$ is purely non-deterministic as is easily seen from the definition, and it is expressed as

$$X_N(t) = \int e^{it\lambda} dZ_N(\lambda)$$

with a Gaussian random measure $Z_N(\cdot)$. Hence

$$X^{[1]}(t) = e^{-(2N-3)t} \frac{d}{dt} \int e^{(i\lambda+2N-3)t} dZ_N(\lambda)$$

exists and is

$$\int (i\lambda + 2N-3) e^{it\lambda} dZ_N(\lambda) \ .$$

The covariance function of $X^{[1]}(t)$ is

$$\int e^{ih\lambda} \{\lambda^2 + (2N-3)^2\} \, |\hat{F}^N(\lambda)|^2 d\lambda$$

$$= -\gamma_N''(h) + (2N-3)^2\gamma_N(h) = c_N^2 \gamma_{N-2}(h) \ ,$$

where $|F_N(\lambda)|^2$ is the spectral density function of $X_N(t)$. Hence $X_N^{[1]}(t)$ can be regarded as a version of $c_N X_{N-2}(t)$.

Thus we can see by (III. 9) and (III, 10) that the study of $X_N(t)$ is reduced to the study of $X_2(t)$ or $X_3(t)$ so far as we consider. More exactly, if we denote the operator $c_N^{-1}e^{-(2N-3)t}\dfrac{d}{dt}e^{(2N-3)t}$ by D_N, it is easily proved that D_3 can be operated to $X_3(t)$ and

$$D_3 D_5 \cdots D_{2p+1} X_{2p+1}(t) = X_1(t) ,$$

which is the Ornstein-Uhlenbeck's Brownian motion. Hence $X_{2p+1}(t)$ and therefore $M_{2p+1}(t)$ is a $(p+1)$-ple Markov process in the restricted sense. (This fact was proved by Lévy by another method). The spectral distribution function of $X_{2p+1}(t)$ has a density

(III. 11) $$|\hat{F}_{2p+1}(\lambda)|^2 = \frac{\prod\limits_{k=1}^{p} c_{2k+1}}{\{(4p-1)^2+\lambda^2\}\{(4p-5)^2+\lambda^2\} \cdots \{3^2+\lambda^2\}} \cdot \frac{1}{1+\lambda^2}$$

Therefore we obtain the following

$$X_{2p+1}(t) = \int_{-\infty}^{t} \left(\sum_{k=1}^{p} a_k e^{-(4k-1)(t-u)} + a_0 e^{-(t-u)}\right) dB_0(u)$$

Here the kernel of the representation is to be determined so that the square of its Fourier transform is equal to $|\hat{F}_{2p+1}(\lambda)|^2$ and it satisfies the condition that $Q(\lambda) =$ constant. This is posible.

Changing the time scale, we have the canonical representation of $M_N(t)$.

$$M_N(t) = \int_{0}^{t} \frac{1}{\sqrt{2}} (\sum_{k=1}^{p} a_k(u/t)^{2k-1} + a_0) dB_0(u) .$$

Obviously thus obtained representation coincides with the Lévy's result.

The problem to obtain all non canonical representations of $M_{2p+1}(t)$, where kernels are polynomials of (u/t) of degree $2p-1$ is easily solved, if we observe the spectral density function of $X_{2p+1}(t)$. The answer of the problem is that *"the number of different representations of above stated form of $M_{2p+1}(t)$ is just 2^{p-1} including canonical one"*.

Proof. If a kernel is a polynomial of (u/t) of degree $2p-1$, it turns into the sum of exponential functions such as $e^{-(2k+1)(t-u)}$, $k \leq 2p-1$, the Fourier transform of which is a sum of the functions of the form $\dfrac{1}{i\lambda + (2k+1)}$. Hence the number of posible functions of $\hat{F}_{2p+1}(\lambda)$ is just 2^{p-1}. For example we obtain such a function

multiplying $\dfrac{i\lambda - (4p - 7)}{i\lambda + (4p - 7)}$ to the function $\hat{F}_{2p+1}(\lambda)$ which corresponds to the canonical kernel.

Example III.3. One of the non-canonical representation of $M_7(t)$ different from the Lévy's one (cf. P. Lévy [4] p. 146) is given as follows: let

$$\hat{F}_7(\lambda) = \left(\frac{c}{(i\lambda + 1)(i\lambda + 3)(i\lambda + 7)(i\lambda + 11)} \right) \cdot \frac{i\lambda - 5}{i\lambda + 5}$$

The rational function in the bracket corresponds to the canonical kernel. Then $M_7(t)$ is expressed as

$$\int_0^t (3/5 - 3u/t + 5u^2/t^2 - 3u^3/t^3 + 2u^5/5t^5)\sqrt{10} dB_0(u) \,.$$

If N is even, we also have

$$D_4 D_6 \cdots D_{2p} X_{2p}(t) = X_2(t) \,.$$

Hence, if we know the canonical kernel of the representation of $X_2(t)$, then we can obtain that of $X_{2p}(t)$ easily and know the properties of it.

We have

(III. 12) $\quad |\hat{F}_{2p}(\lambda)|^2 = \dfrac{\prod\limits_{k=1}^{p} c_{2k}}{\{\lambda^2 + (4p-3)^2\} \{\lambda^2 + (4p-7)^2\} \cdots \{\lambda^2 + 5^2\}} |\hat{F}_2(\lambda)|^2$

in the way similar to the case in which N is odd. Now it is our purpose to obtain the exact form of $|\hat{F}_2(\lambda)|^2$. To do so, let us consider $\gamma_2(h)$. If $h > 0$,

$$\gamma_2(h) = \cosh h/2 - \frac{1}{2\pi} \int_0^\pi (2\cosh(2h) - 2\cos\theta)^{1/2} d\theta$$

Using the Legendre's polynomials the integral term of it may be expanded as follows

$$\frac{1}{\sqrt{2}\pi} \int_0^\pi \{\cosh(2h) - \cos\theta\}^{1/2} d\theta = \frac{1}{2} \{e^h + \sum_{k=0}^\infty (a_{k+1}^2 + a_k^2 - 2a_k a_{k+1})e^{-(4k+3)h}\} \,,$$

where

$$a_k = \frac{1 \cdot 3 \cdot 5 \cdots (2k-1)}{2 \cdot 4 \cdot 6 \cdots 2k.}$$

Taking the Fourier transform, we have

$$|\hat{F}_2(\lambda)|^2 = \frac{1}{2\pi} \left(\frac{1}{1+\lambda^2} - \sum_{k=0}^\infty \frac{b_k}{\lambda^2 + (4k+3)^2} \right),$$

where

$$b_k = (4k+3)(a_{k+1} - a_k)^2 \qquad (>0).$$

This proves that $X_2(t)$ is not a multiple Markov process, but, as it were, ∞-ple Markov process.

Let $\hat{F}^{(n)}(\lambda)$ be given by

$$\hat{F}^{(n)}(\lambda) = \left(\frac{1}{1+\lambda^2} - \sum_{k=0}^{n-1} \frac{b_k}{\lambda^2 + (4k+3)^2} \right) / 2\pi .$$

Then $\hat{F}^{(n)}(\lambda) > 0$ (since $\hat{F}^{(n)}(\lambda) > |\hat{F}_2(\lambda)|^2 \geqq 0$) and

$$\int \frac{\log \hat{F}^{(n)}(\lambda)}{1+\lambda^2} d\lambda > -\infty .$$

Therefore, there exists a stationary Gaussian process $X^{(n)}(t)$ which is expressed in the form

$$X^{(n)}(t) = \int_{-\infty}^{t} F_n(t-u) dB_0(u)$$

and has spectral density function $\hat{F}^{(n)}(\lambda)$.

Obviously $X^{(n)}(t)$ is a stationary $(n+1)$-ple Markov process and its covariance function $\gamma^{(n)}(h)$ converges to $\gamma_2(h)$ uniformly in any finite interval of h.

Summing up we have

Theorem III. 2. *$X_{2p}(t)$ is not a multiple Markov process, but it is a limiting process of n-ple Markov process $(n \to \infty)$.*

Proof. For $p=1$, we have already proved. Noting the formula (III. 12), we can easily prove our theorem.

From this theorem we can see that $M_{2p}(t)$ is not a multiple Markov process, but is a limiting process of n-ple Markov process with homogeneous canonical kernel of degree 0, which was to be obtained.

REFERENCES

N. Aronszajn [1]: *Theory of reproducing kernels.* Trans. Amer. Math. Soc. vol. 68, 337–404 (1950).

J. L. Doob   [1]: *The elementary Gaussian processes.* Annals of Math. Stat. vol. 15, 229–282 (1944).

            [2]: *Stochastic processes.* Willey, (1952).

C. L. Dolph and M. A. Woodbury
            [1]: *On the relation between Green's functions and covariances of certain stochastic processes and its application to unbiased linear prediction.* Trans. Amer. Math. Soc. vol. 72, 519–550 (1952).

E. L. Ince   [1]:   *Ordinary differential equations.* Dover Pub. INC (1926).

K. Itô   [1]:   *Stochastic integral.* Proc. Imp. Acad. Tokyo, vol. 20, 519–524 (1944).

       [2]:   *Theory of Probability.* (In Japanese). Tokyo, (1953).

       [3]:   *Stochastic process.* Tata Institute Note, Bombay, (1959).

S. Itô   [1]:   *On Hellinger-Hahn's Theorem.* (In Japanese). Sūgaku, vol. 5 no. 2, 90–91 (1953).

K. Karhunen   [1]:   *Über die Struktur stationärer zufälliger Funktionen.* Arkiv för Matematik, 1. nr. 13, 144–160 (1950).

P. Lévy   [1]:   *Processus stochastiques et mouvement Brownien.* Gautier-Villars, Paris, (1948).

       [2]:   *Le mouvement brownien.* Memorial des Science Math. fasc. 129, (1954)

       [3]:   *Brownian motion depending on n parameters: The particular case* $n=5$. Proceedings of Symposia in Applied Math. vol. 7, 1–20 (1957).

       [4]:   *A special problem of Brownian motion, and a general theory of Gaussian random functions.* Proceedings of the Third Berkeley Symposium on Math. Stat. and Prob. vol. II. 133–175 (1956).

       [5]:   *Sur une classe de courbes de l'espace de Hilbert et sur une équation intégrale non linéaire.* Annales École Norm. Sup. tom. 73, 121–156 (1956).

       [6]:   *Fonction aléatoires à corrélation linéaire.* Illinois Journal of Math. vol. 1, 217–258 (1957).

       [7]:   *Fonctions linéairement Markoviennes d'ordre n.* Math. Japonicae, vol. 4, no. 3, 113–121 (1957).

       [8]:   *Sur quelques chasses de fonctions aléatoires.* Journal de Math. pures et appliquées. tom. 38, 1–23 (1959).

L. Schwartz   [1]:   *Théorie des distribution I.* Hermann & Cie, Paris, (1950).

T. Seguchi and N. Ikeda

       [1]:   *Note on the statistical inferences of certain continuous stochastic processes.* Memoires of the Faculty of Sci. Kyūsyū Univ. Ser. A, vol. 8, No. 2, 187–199 (1954).

T. Sirao   [1]:   *On the continuity of Brownian motion with a multidimensional parameter.* Nagoya Math. Journal, Vol. 16, 135–156 (1960).

M. H. Stone   [1]:   *Linear transformations in Hilbert space and its applications to analysis.* Amer. Math. Soc. Colloq. Pub. vol. 15. (1932).